

# Data-driven weather forecasting

Mariana Clare

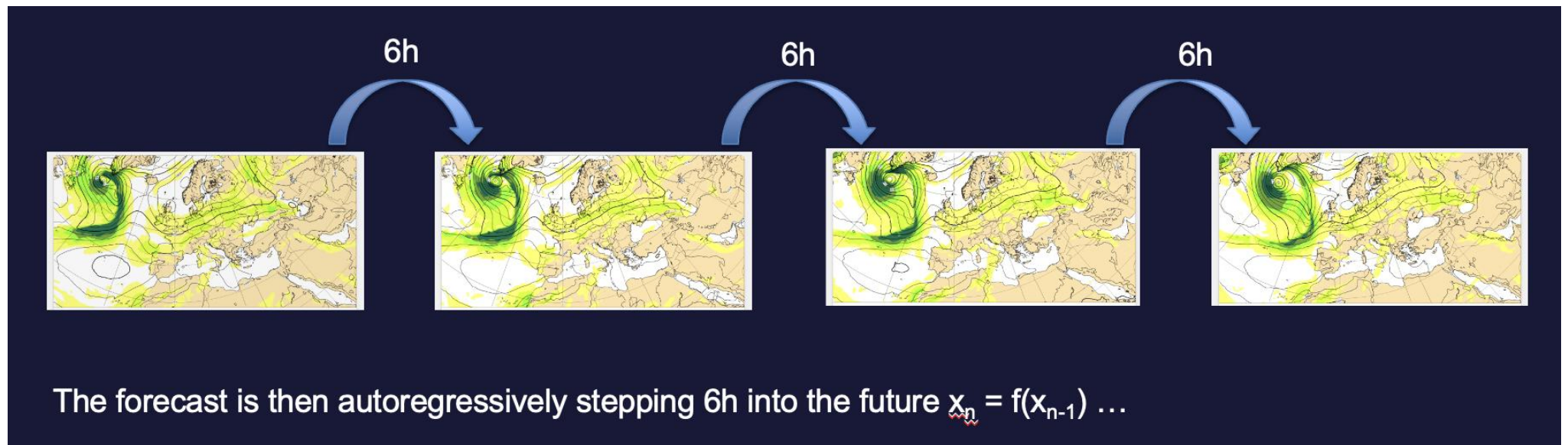
# Classical NWP



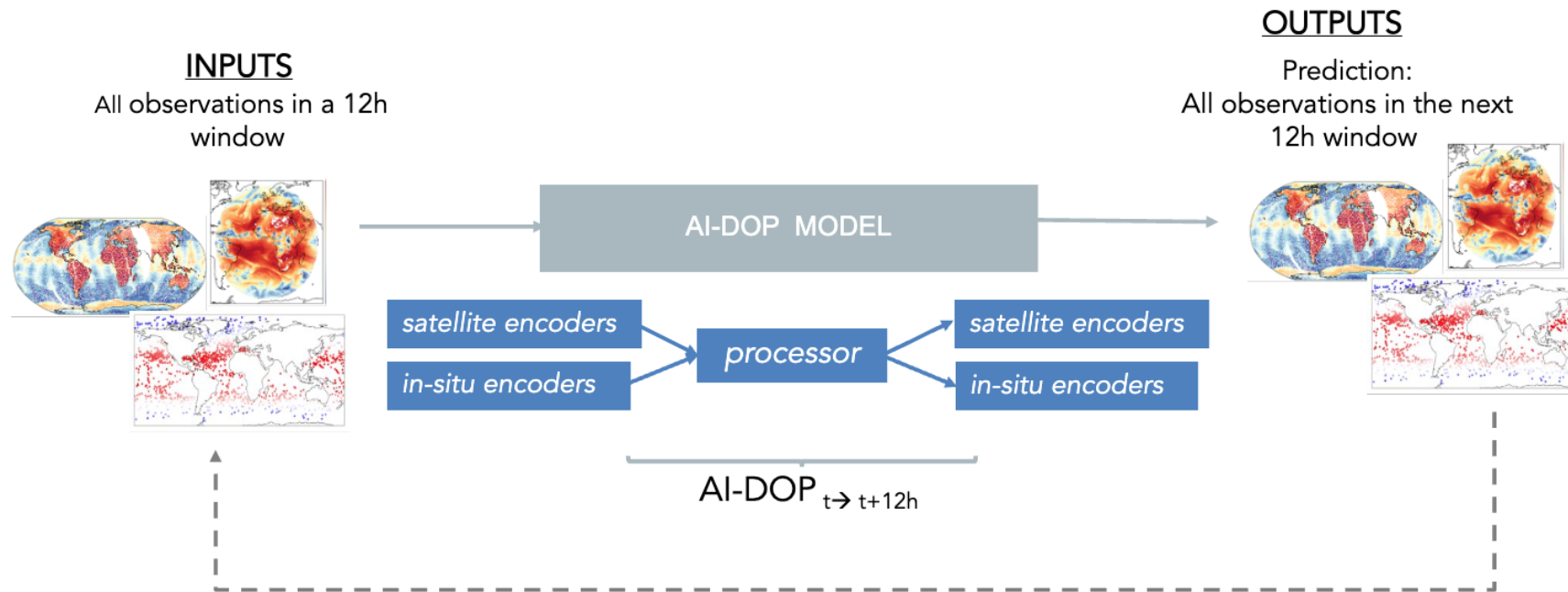
# Data-driven weather forecasting system



Most data-driven weather forecasting models are trained on reanalysis datasets



# Data-driven weather forecasting system



*Training on observations (e.g. SYNOPs) at arbitrary locations allows for generalisation to standard gridded model output*

# A short history of data-driven weather forecasting

February 2022 – First competitive medium-range systems

- **Keisler** – GraphNN, competitive with GFS (USA)
- **NVIDIA** – **FourCastNet Fourier+**, 0.25°, **O(10<sup>4</sup>) faster & more energy efficient than IFS**

December 2022

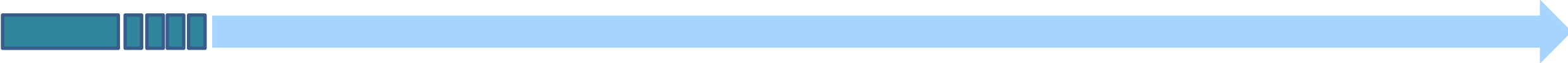
- **Deepmind** – **GraphCast**

GraphNN  
0.25° **Many parameters with comparable skill to IFS.**

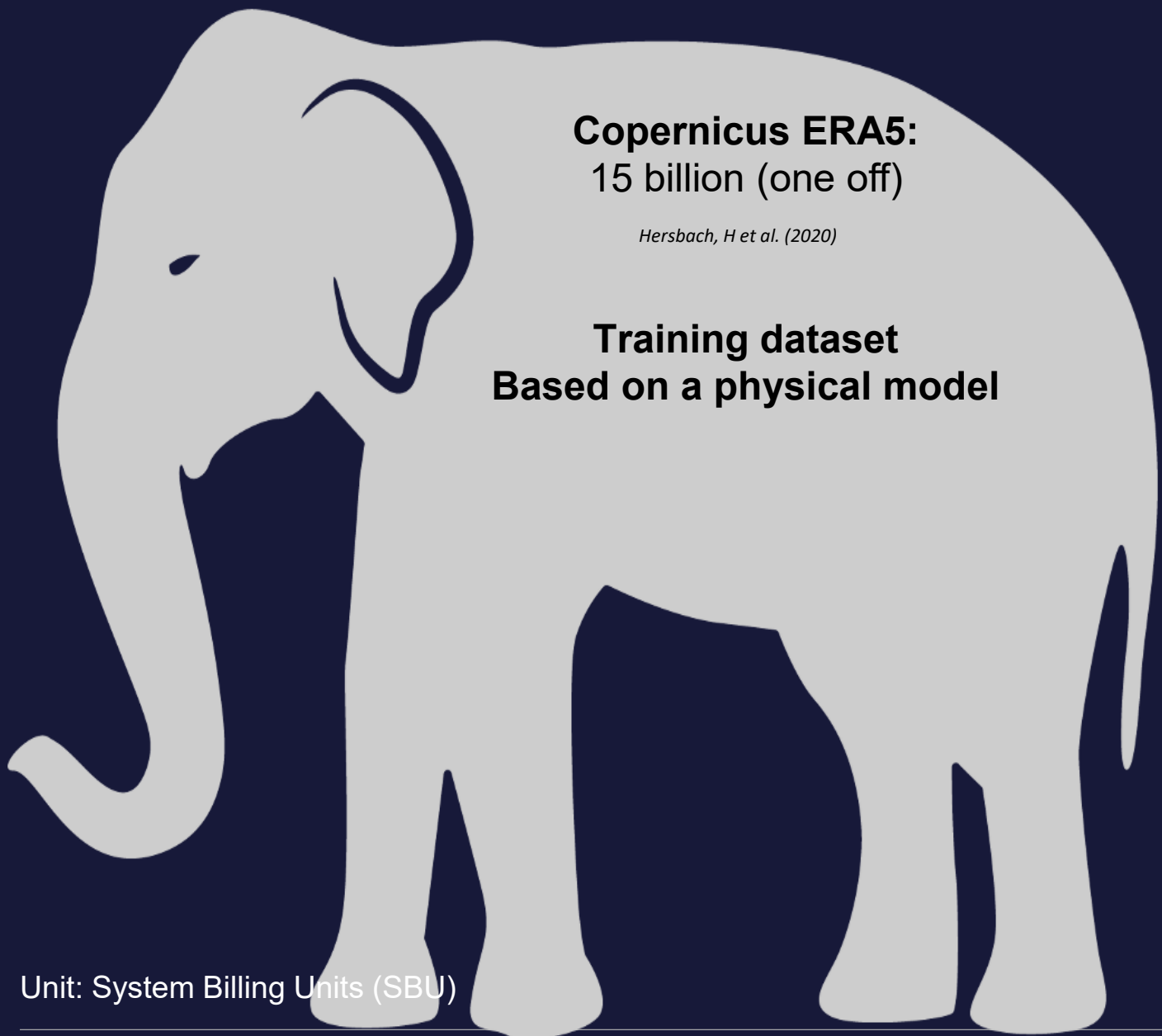
November 2022

- **Huawei** – **PanguWeather**

Vision Transformer  
0.25° **“More accurate tropical cyclone tracks” than the IFS.**



2018 – Concept explored (ECMWF and others)...



**Copernicus ERA5:**  
15 billion (one off)

*Hersbach, H et al. (2020)*

**Training dataset  
Based on a physical model**

**Physics-based Forecast:**

180 000  
per forecast

**AI Model:**

0.3  
per forecast

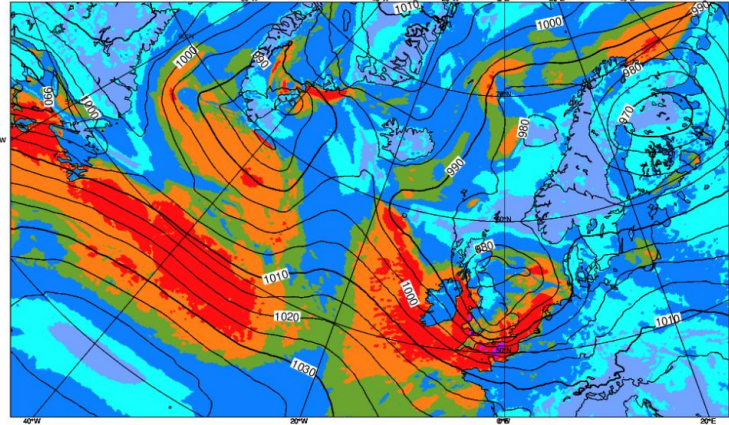
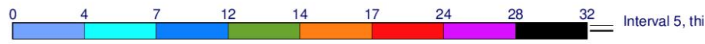
~1000x  
reduction in  
energy  
Forecast time  
reduced from  
~30minutes to  
~3minutes



Unit: System Billing Units (SBU)

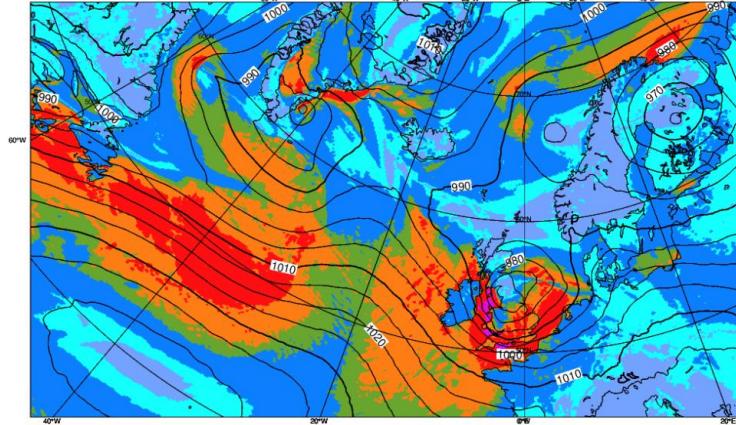
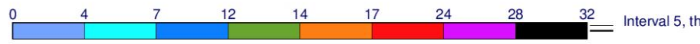
## Analysis

MSLP+WS 2022021612 Step: 0  
AN



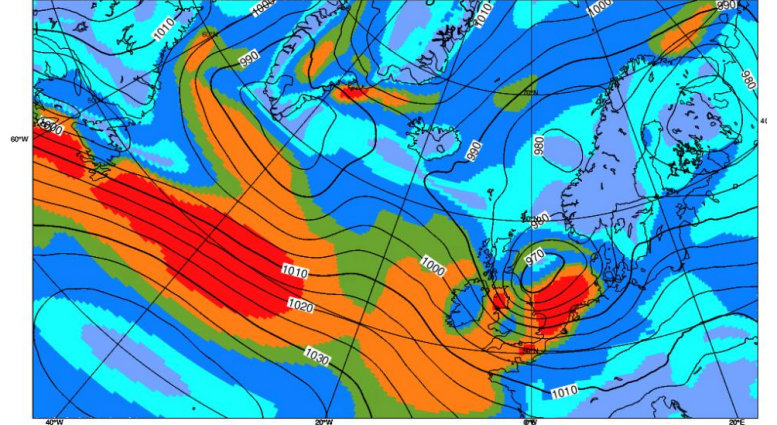
## HRES

MSLP+WS 2022021600 Step: 60  
HRES



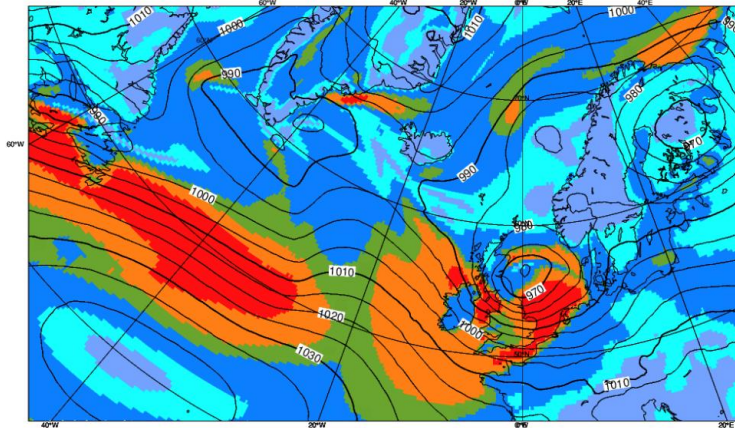
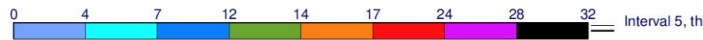
## FourCastnet

MSLP+WS 2022021600 Step: 60  
i51c



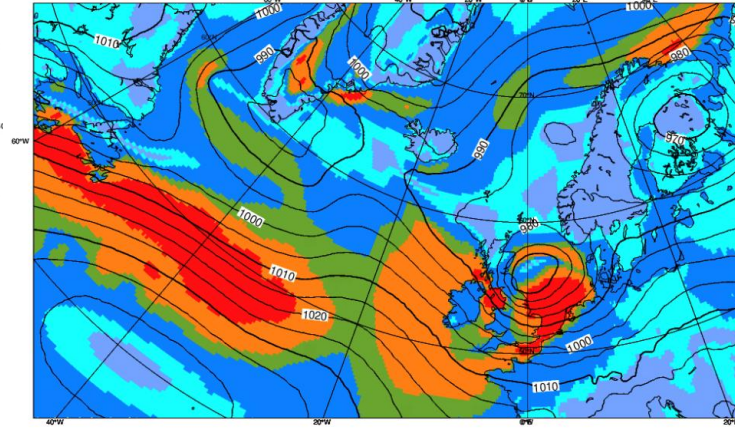
## PanguWeather

MSLP+WS 2022021600 Step: 60  
PanguWeather



## GraphCast

MSLP+WS 2022021600 Step: 60  
i51d



# A short history of data-driven weather forecasting

February 2022 – First competitive medium-range systems

- **Keisler** – GraphNN, competitive with GFS (USA)
- **NVIDIA** – FourCastNet Fourier+ , 0.25°, O(10<sup>4</sup>) faster & more energy efficient than IFS

December 2022

- **Deepmind** – GraphCast

GraphNN  
0.25° Many parameters with comparable skill to IFS.

November 2022

- **Huawei** – PanguWeather

Vision Transformer  
0.25° “More accurate tropical cyclone tracks” than the IFS.

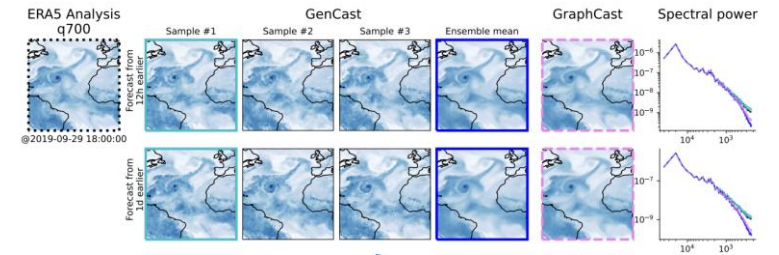
January-June 2023

- **Microsoft** – ClimaX
- **China academia/Shanghai Met** – FengWu
- **Alibaba** – SwinRDM
- **NVIDIA** – SFNO
- ...

December 2023

- **Deepmind** – GenCast

Probabilistic forecast (ensemble) – 1°  
”Outperforming the leading operational ensemble forecast” (aka ECMWF)

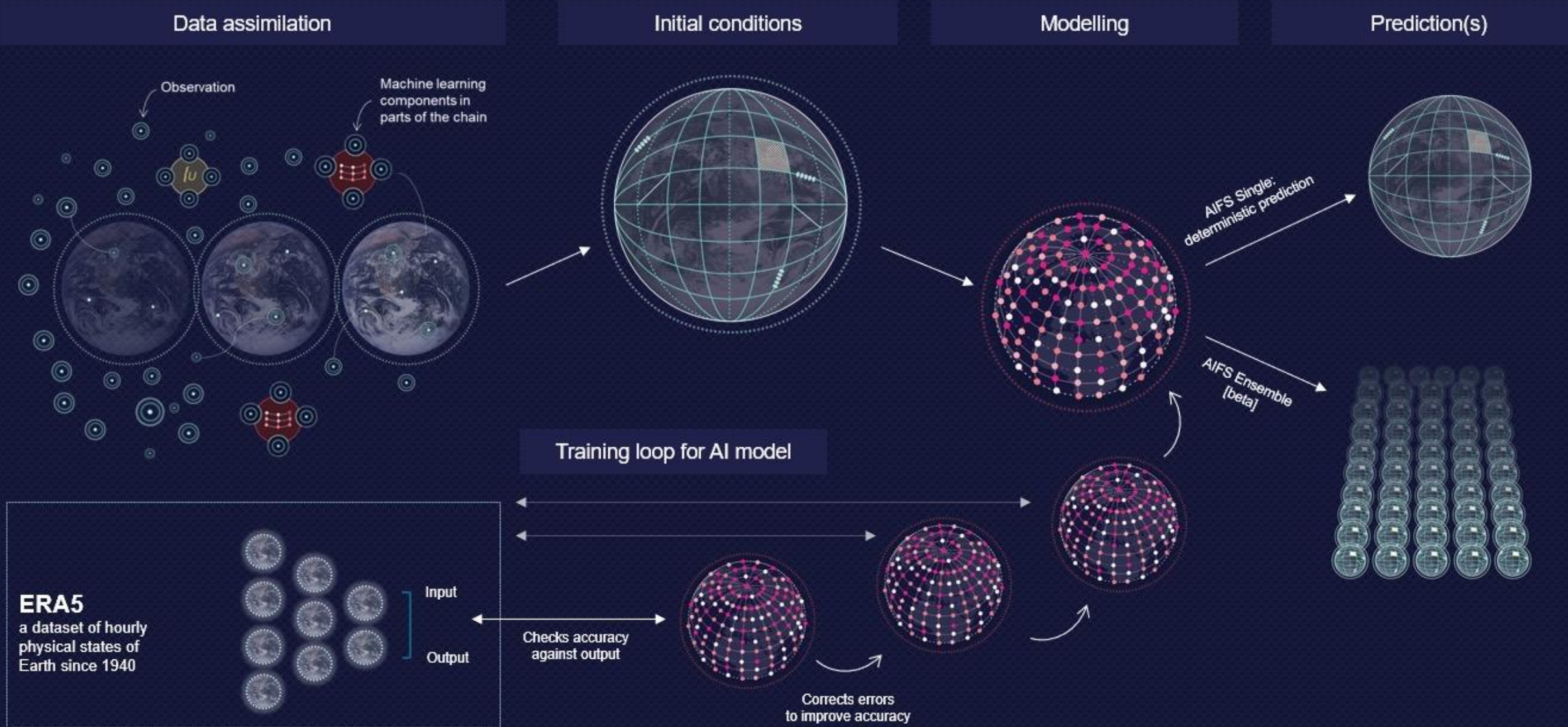


2018 – Concept explored (ECMWF and others)...

June 2023  
ECMWF – ML project begins

Early 2023  
Prototype AIFS developments begin

# AIFS: Artificial Intelligence Forecasting System



# A short history of data-driven weather forecasting

**February 2022 – First competitive medium-range systems**

- **Keisler** – GraphNN, competitive with GFS (USA)
- **NVIDIA** – FourCastNet Fourier+ , 0.25°, O(10<sup>4</sup>) faster & more energy efficient than IFS

**December 2022**

- **Deepmind** – GraphCast

GraphNN  
0.25° Many parameters with comparable skill to IFS.

**November 2022**

- **Huawei** – PanguWeather

Vision Transformer  
0.25° “More accurate tropical cyclone tracks” than the IFS.

**January-June 2023**

- **Microsoft** – ClimaX
- **China academia/Shanghai Met** – FengWu
- **Alibaba** – SwinRDM
- **NVIDIA** – SFNO
- ...

**December 2023**

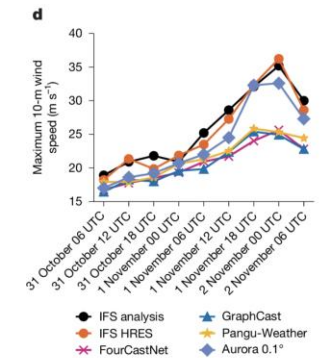
- **Deepmind** – GenCast

Probabilistic forecast (ensemble) – 0.25°  
”Outperforming the leading operational ensemble forecast” (aka ECMWF)

**June 2024**

- **Microsoft** – Aurora

Higher resolution – 0.1°  
Atmospheric composition



**2018 – Concept explored (ECMWF and others)...**

**June 2023  
ML project begins**

**Jan/Feb 2024  
AIFS first updates**

**July 2025:  
AIFS ENS CRPS  
1 operational**

**Early 2023  
Prototype AIFS  
developments  
begin**

**October 2023  
AIFS experimental  
forecasts live**

**Feb 2025:  
AIFS Single 1  
operational**

# Different model architectures

# Convolutions

- Simplest design, treat the Earth as a cylinder, use convolutions with periodicity in longitude.
  - No treatment of the pole. How can flow easily pass over the artic?
- Weyn et al 2020 proposed a clever cubed-sphere approach.
  - One set of CNN for the side faces of the cube.
  - Another for the polar faces.
- Karlbauer et al (2023) do convolutions on the HealPix grid (see below)

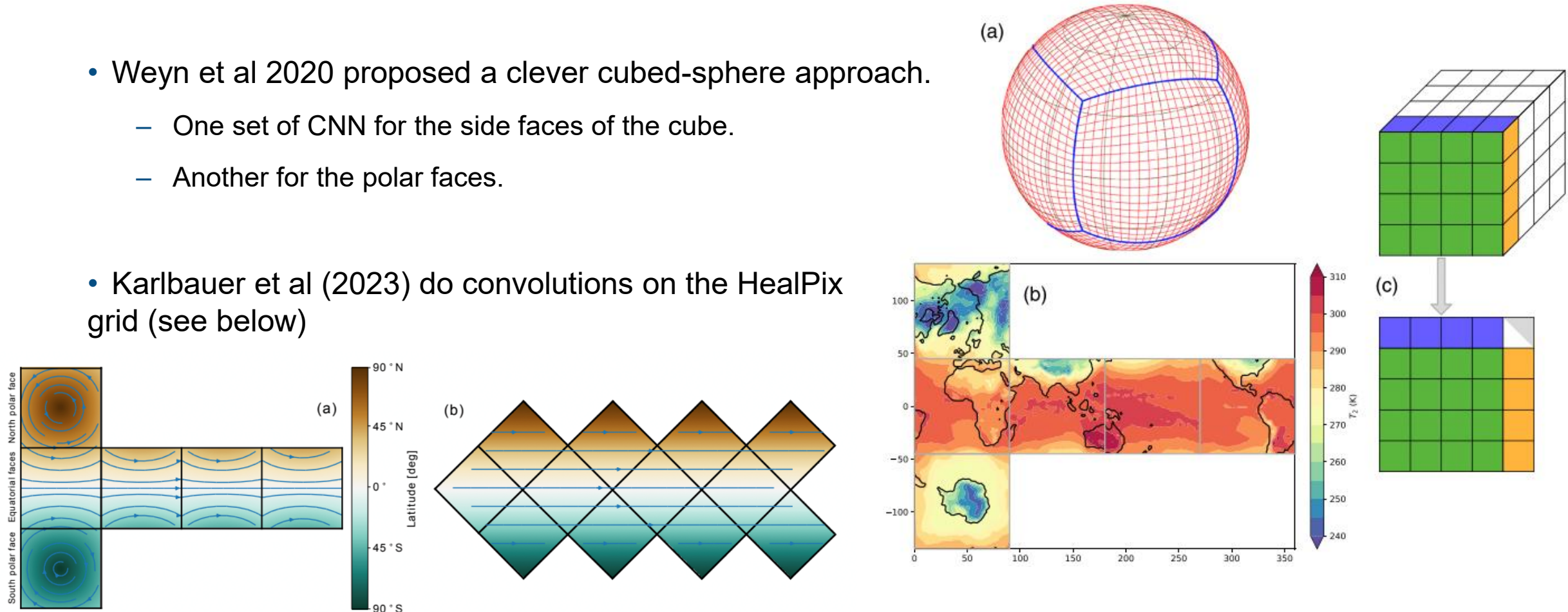
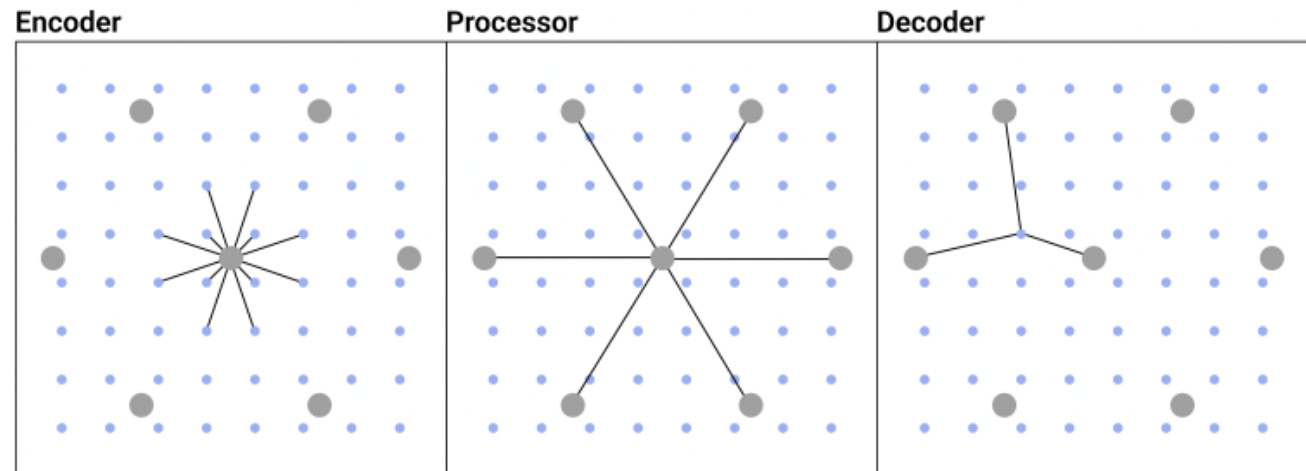
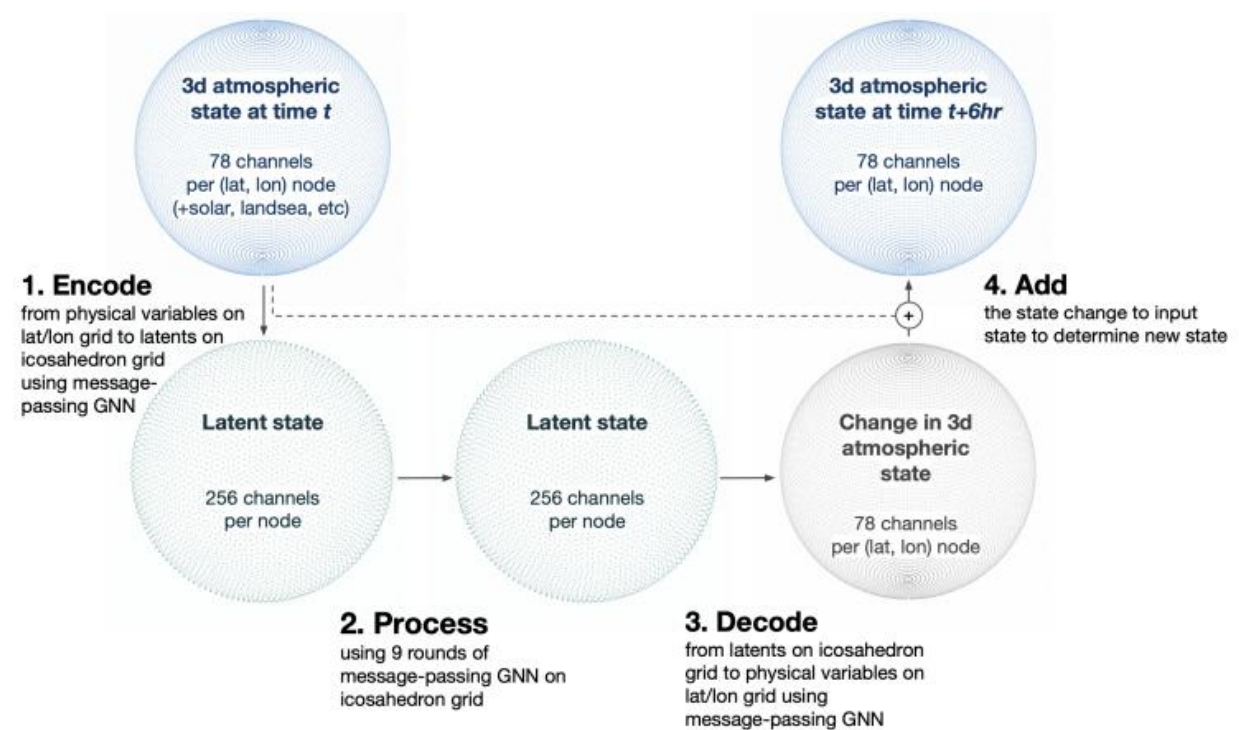


Figure 10: Lines of latitudes depicted as blue streamline arrows on the cubed sphere (a) and on the HEALPix (b). While the lines corresponding to constant eastward motion describe arcs of different radii on the cubed sphere mesh, the same motion translates to straight lines on the HEALPix mesh.

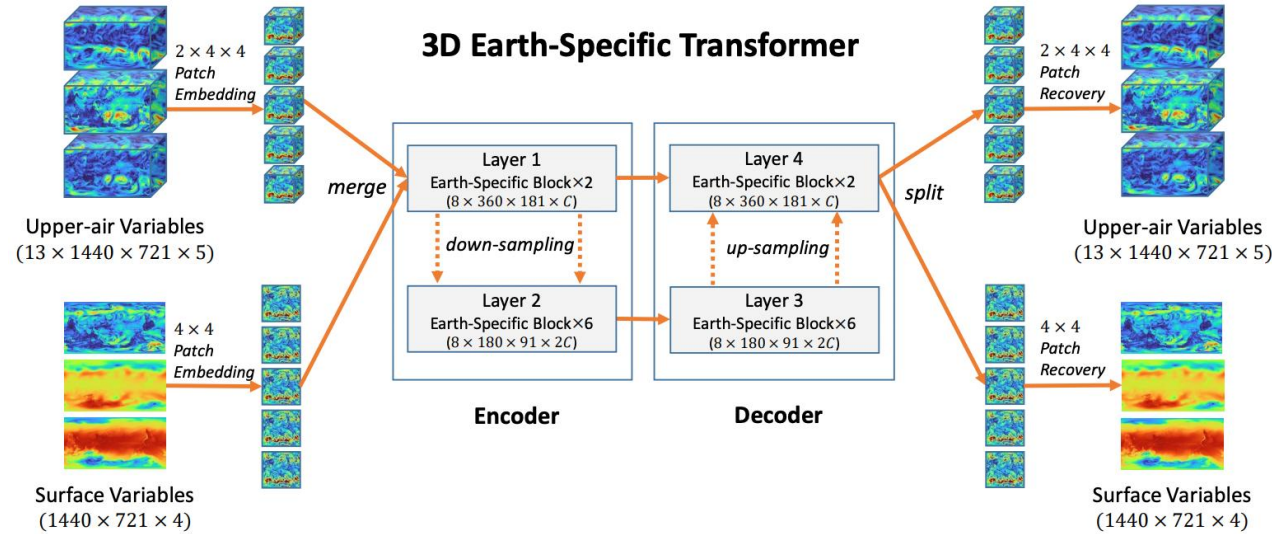
# Graph Neural Networks

- First demonstrated by Keisler 2022.
- Data is structured in nodes and the connections between them (edges)
- Most popular, the message passing GNN.
  - Involves MLPs on the edges, and nodes.
  - Alternates are GraphConvolutions & Graph Attention
- Can handle unstructured data in space.

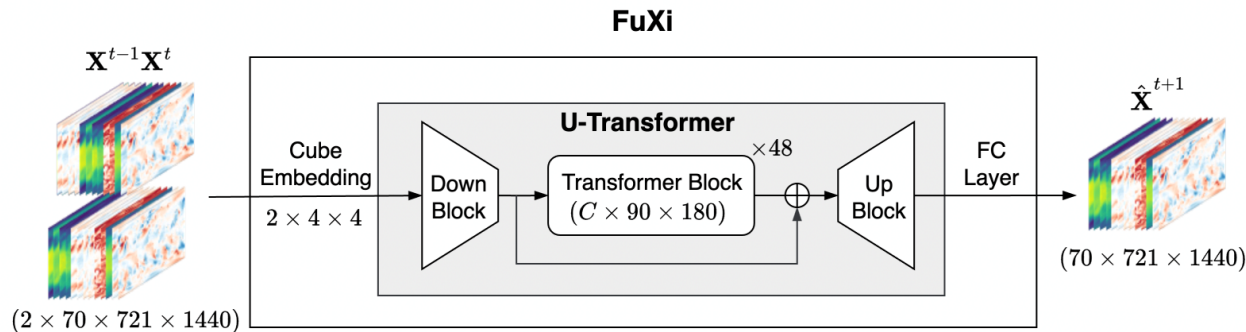


# (Vision) Transformers

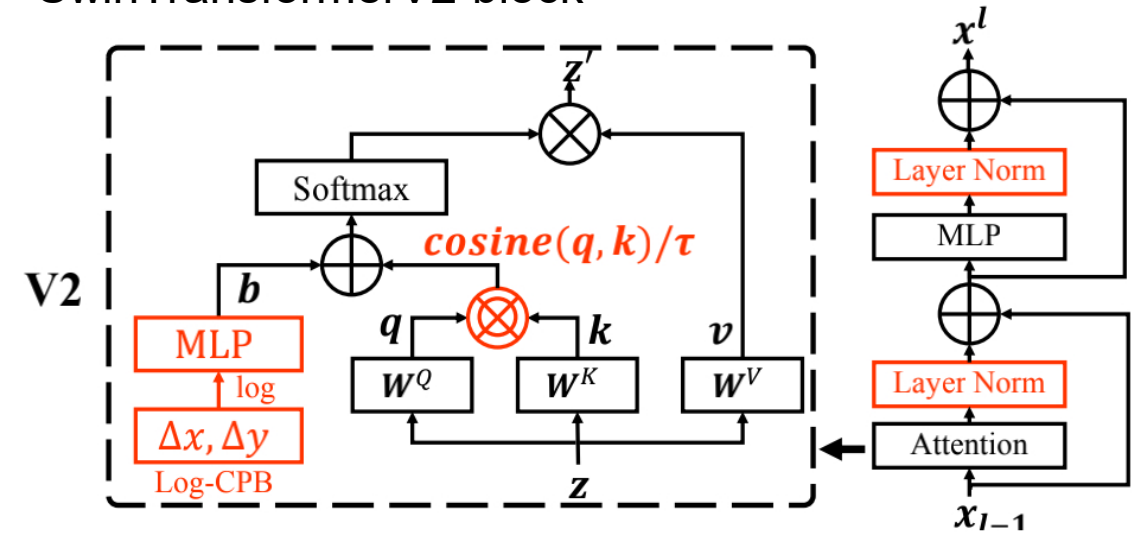
- Build heavily on advances in vision transformers
- Pangu, FuXi, FengWu.
- Embed to a coarser resolution.
- Shifted-window approach adapted to include longitudinal periodicity but poles are not explicitly handled.



a) The overall architecture of FuXi model

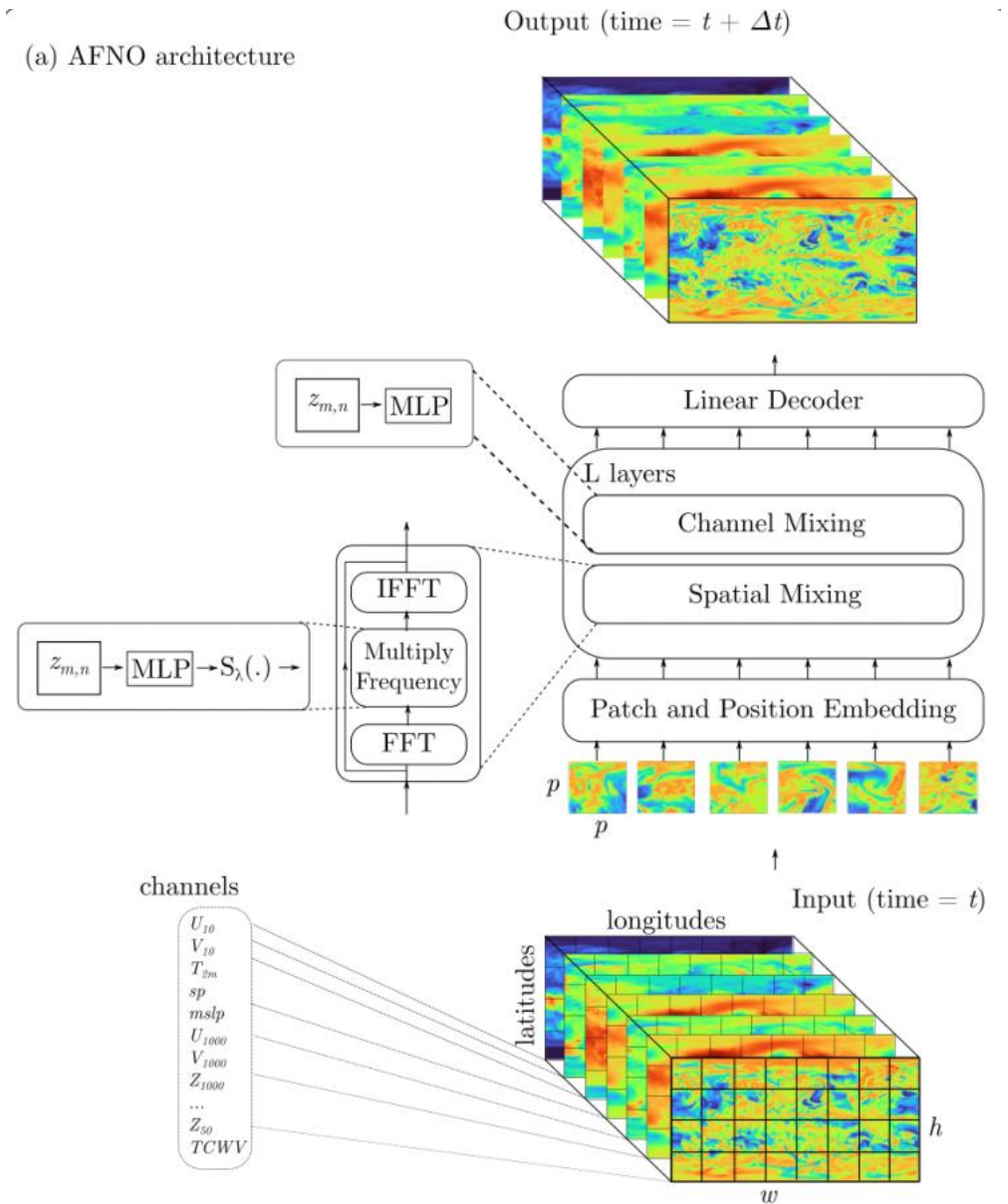


## SwinTransformerV2 block



# Fourier Neural operators

- e.g. FourCastNet, popularised by NVIDIA.
- Part of neural network carried out in frequency space and part in grid-point space.
- Grid-invariance built in.
- Spherical version encodes the symmetries of the sphere.
- Also used in “ACE”, the climate emulator.



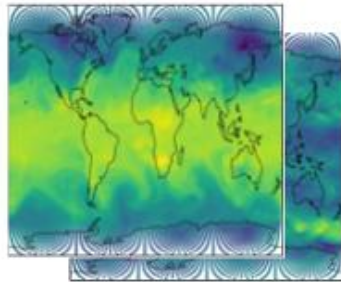
## So, which to choose?

- Vision transformer and GNN solutions both hit comparable levels of skill.
  - SFNO a little further behind in skill, unclear why.
  - CNN not been implemented at the same scale.
- GNN naturally encodes the sphere and allows use of unstructured grids.
- Vision transformers (and SFNO) appear to converge faster than GNN.

# AIFS Single v1

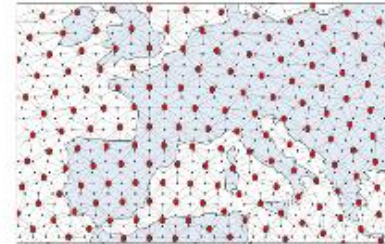
Lang et al 2024a  
Operational system  
from 25/2/25

**INPUTS**  
Atmospheric state:  
 $X(t)$ ,  $X(t-6h)$



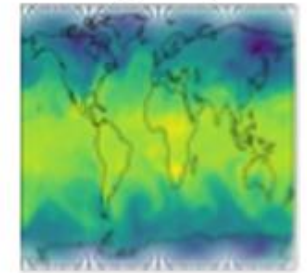
sed)

decoder



**OUTPUTS**

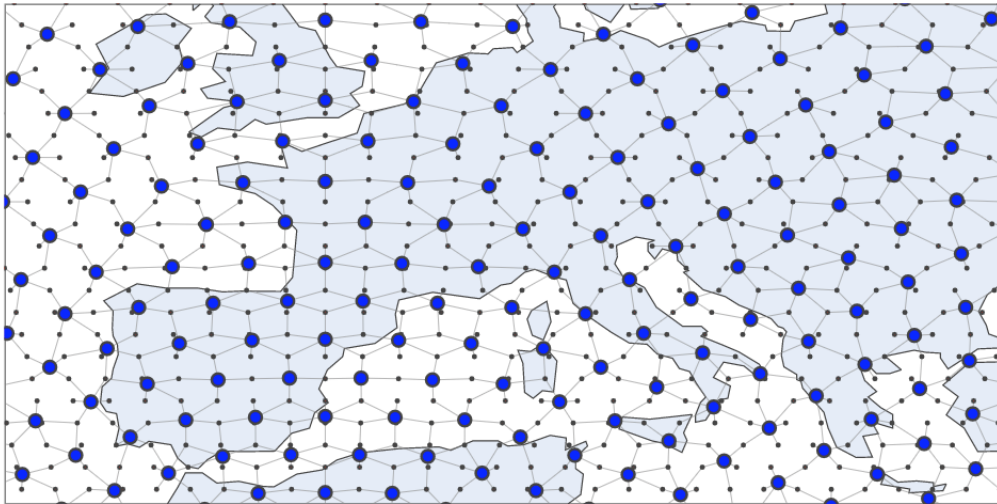
Prediction:  
 $X(t+6h)$



# AIFS – Encoder and Decoder

AIFS works with the native IFS reduced gaussian grids ;  
possible to split model and input data across multiple GPUs to  
handle large memory requirements

Encoder, GNN



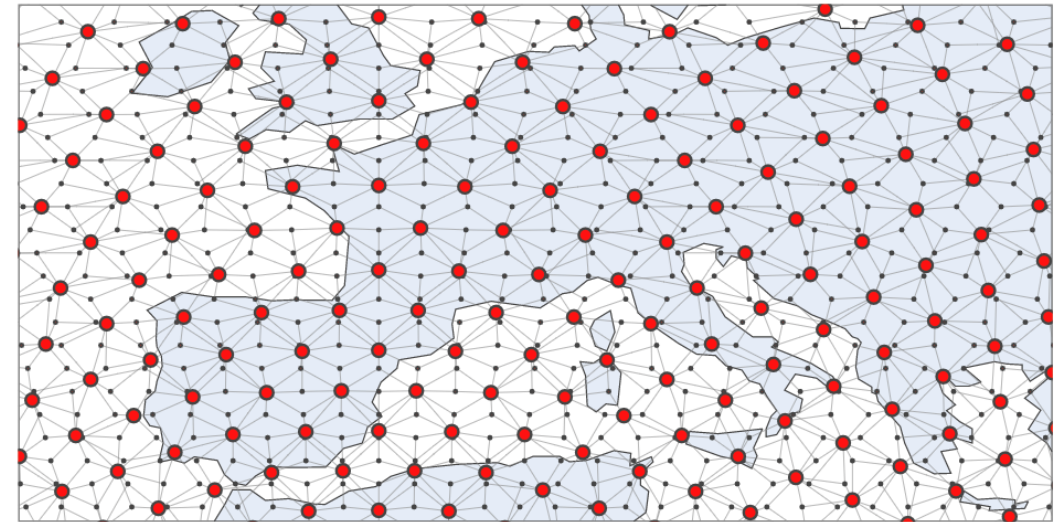
ERA5 – n320

~ 540 000 Nodes,  
~ 1 Million Edges

16 Processor Layers

O96 ~ 40 000 Nodes

Decoder, GNN



ERA5 – n320

# AIFS – Processor

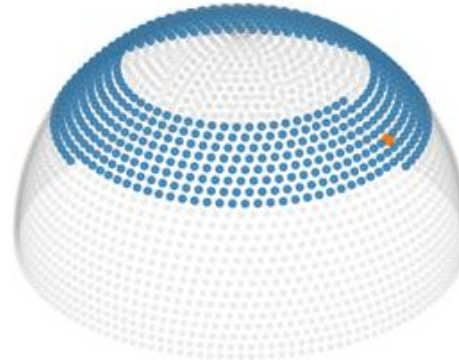
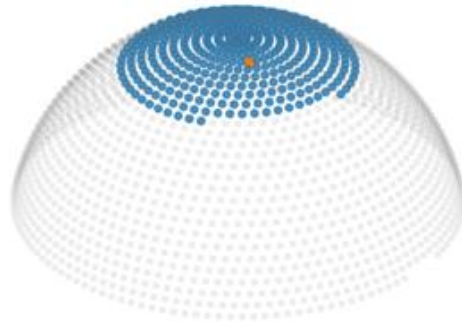
Transformer (like LLMs) that works with a sliding attention window -> attention bands around the globe

Red: target node

Blue: Nodes target node attends to in one processor layer

Grey: How far information can travel within e.g. 6 processor layers

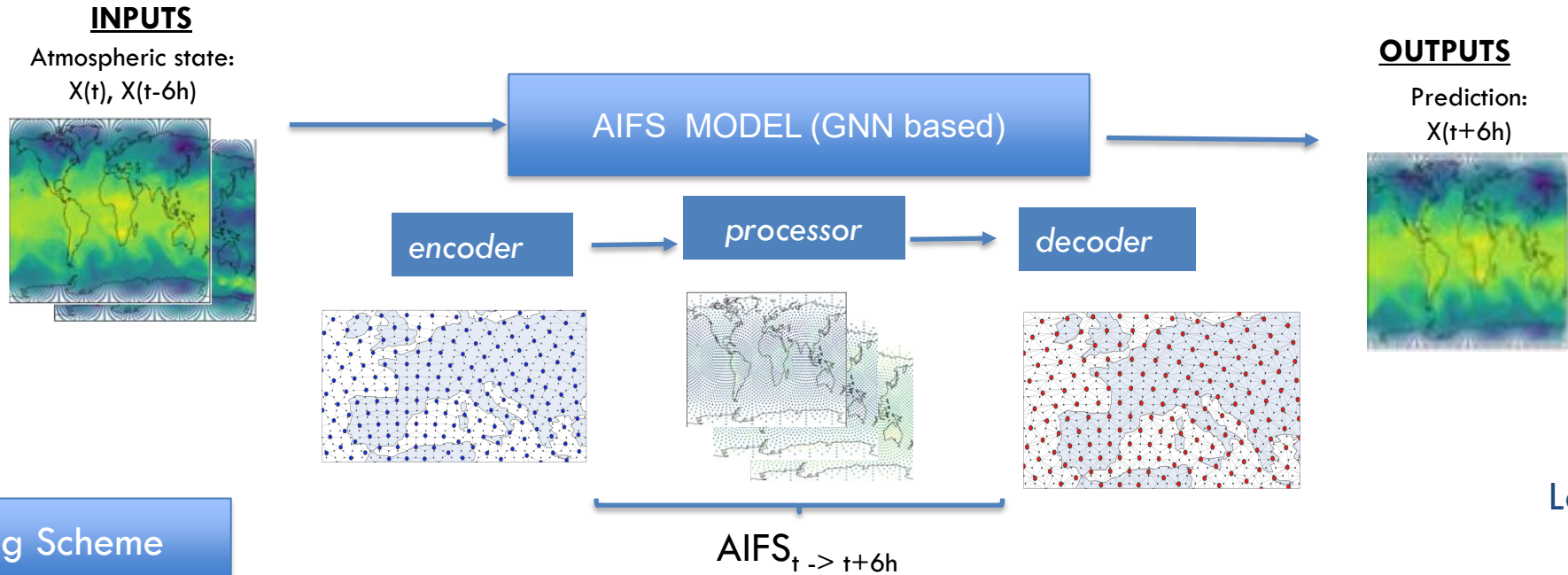
(here lower resolution processor grid than operational AIFS for visualization)



# Training setup

# Training

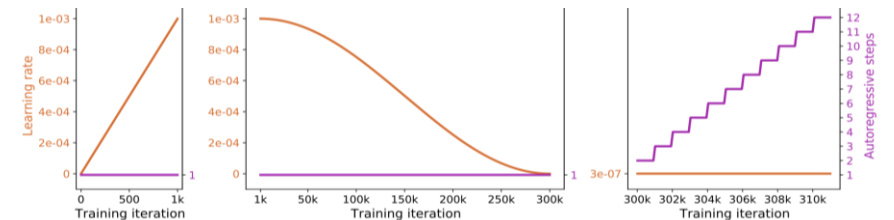
## TRAINING



Lang et al. (2024)

## Training Scheme

- Step 1:** pre-training phase, during which the model is given the task to forecast 6 h ahead
- Step 2:** model is trained to produce forecast up to 72 hr ahead
- Step 3:** model is fine-tuned on operational real-time analysis



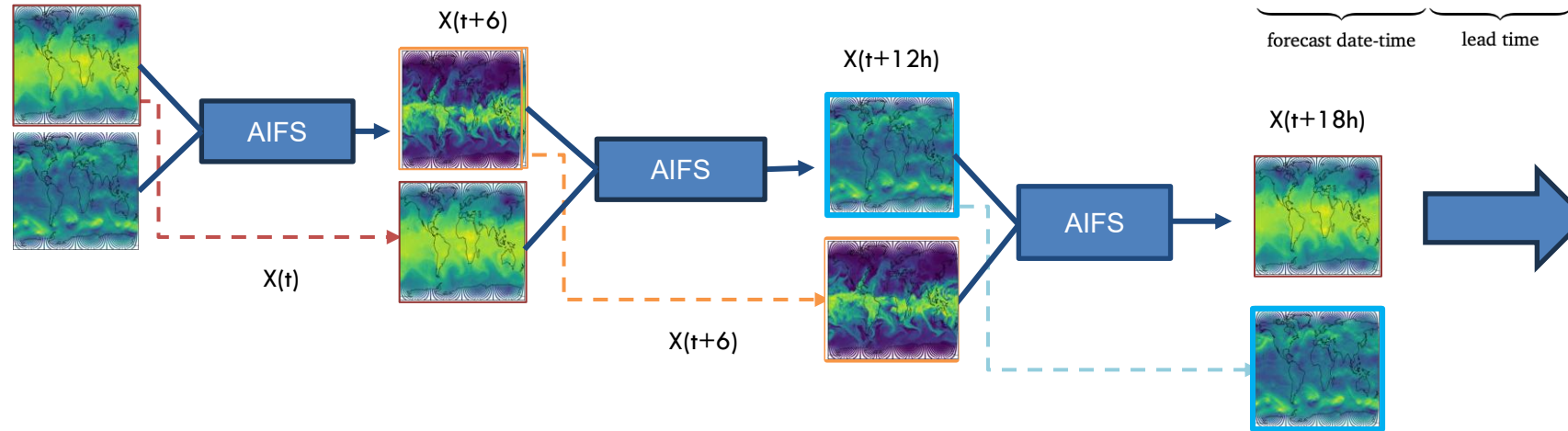
Scheme – Step 1

Scheme – Step 2 & 3

# Training

## TRAINING SCHEME – Step 2

Atmospheric state:  
 $X(t), X(t-6h)$

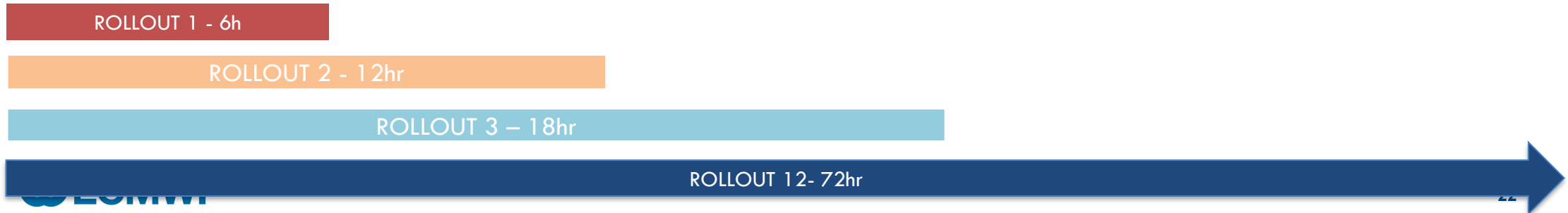


WMSE - area-weighted mean squared error

$$\mathcal{L}_{\text{MSE}} = \underbrace{\frac{1}{|D_{\text{batch}}|}}_{\text{forecast date-time}} \sum_{d_0 \in D_{\text{batch}}} \underbrace{\frac{1}{T_{\text{train}}}}_{\text{lead time}} \sum_{\tau \in 1:T_{\text{train}}} \underbrace{\frac{1}{|G_{0.25^\circ}|}}_{\text{spatial location}} \sum_{i \in G_{0.25^\circ}} \underbrace{\sum_{j \in J}}_{\text{variable-level}} s_j w_j a_i \underbrace{(\hat{x}_{i,j}^{d_0+\tau} - x_{i,j}^{d_0+\tau})^2}_{\text{squared error}}$$

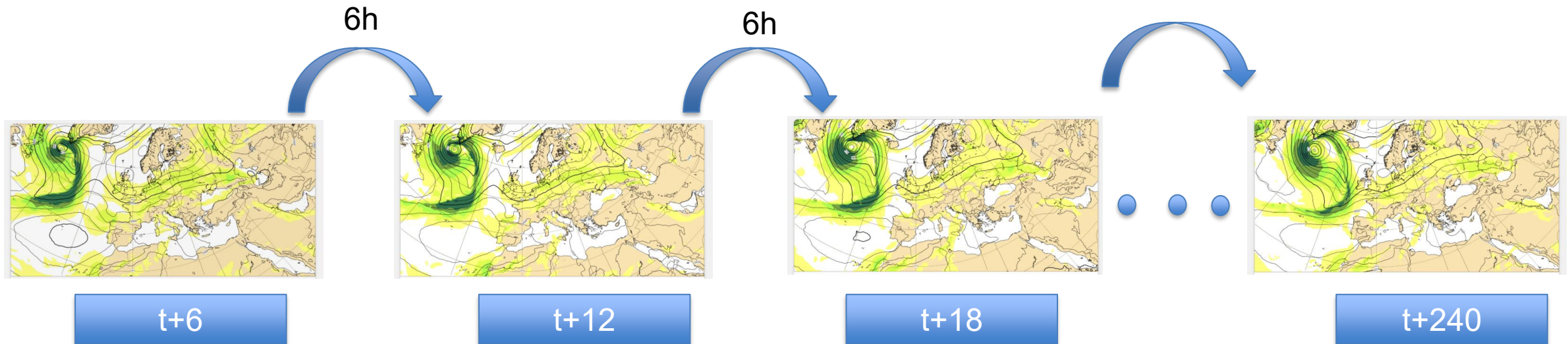


model is trained to produce forecast up to 72 hr ahead by aggregating the WMSE



# AIFS Single v1 - Inference

Once the model is trained, we can generate predictions stepping e.g. 6h from analysis to analysis



The forecast is then autoregressively stepping 6h into the future  $x_n = f(x_{n-1}) \dots$

# What variables to use?

Two driving motivations:

- What helps me predict better.
- What do users want from the system.





Typical set used by many models:

- ~13 pressure levels (with model top at 50hPa)
  - Contrast with 137 model levels in the IFS.
  - GraphCast version with 37 levels isn't more skilful than 13 level version.
- q, t, u, v, w, z
  - But no direct cloud information.
- At the surface: 2t, msl/sp, 10u/v and precipitation for some.






# Variables used in AIFS 2.0

A mix of prognostic, diagnostic, and forcing variables:





## Upper Atmosphere (14 pressure levels)

-  Geopotential height
-  Wind components (u, v, w)
-  Specific humidity
-  Temperature

## Surface Variables

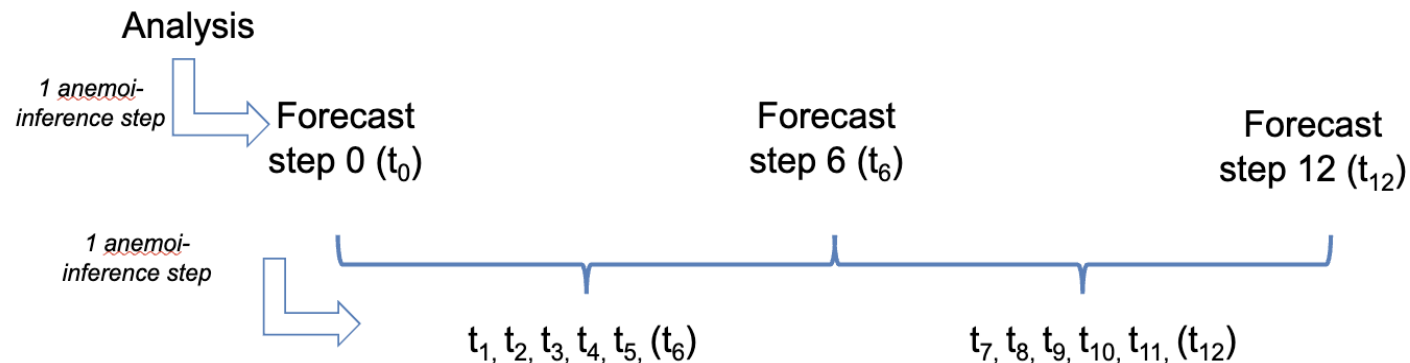
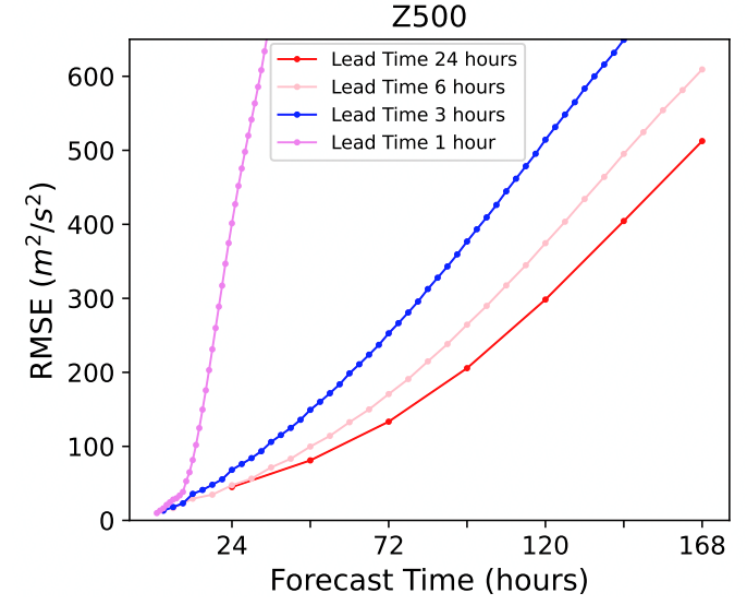
-  2 m temperature
-  10 m wind speed
-  Precipitation (including snow)
-  Clouds
-  Waves

## External Forcings

-  Orography
-  Insolation
-  Latitude / Longitude
-  Time of day / Time of year

# How to handle time?

- How many time slices to provide as input?
- How big of a timestep to make?
  - Early work tried 3-day steps but failed to compete with physics models.
  - Many choose 6-hours, but this limits the granularity of the output.
  - Pangu Weather created multiple models for different timesteps...
  - but this leads to inconsistencies in time.
- Exploring using a temporal downscaler to increase temporal resolution

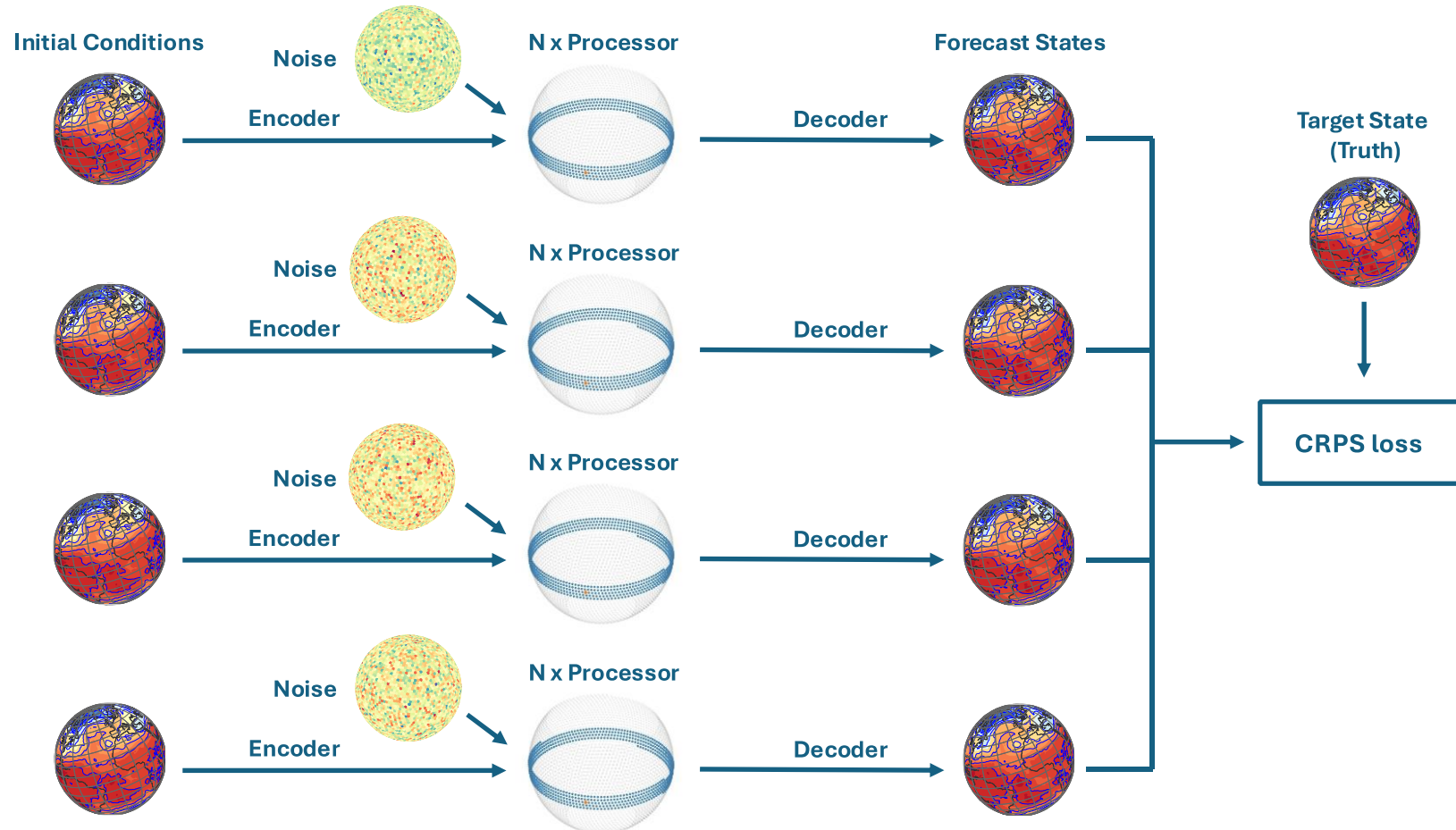


# Proper score loss – AIFS-ENS (AIFS-CRPS):

In training: run (small) ensemble:

Lang et al 2024b

Operational system from 1/7/25



To generate a forecast (single member): run model with noise realization for each forecast step

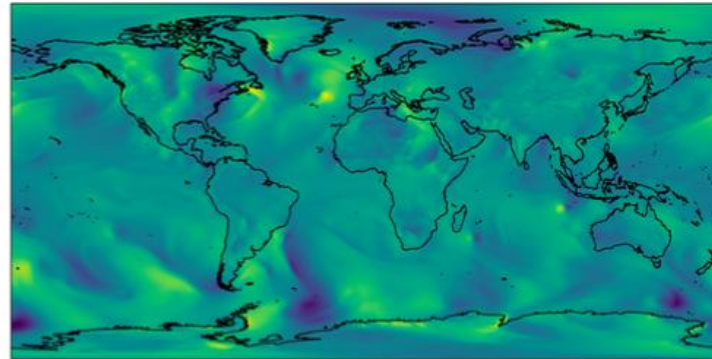
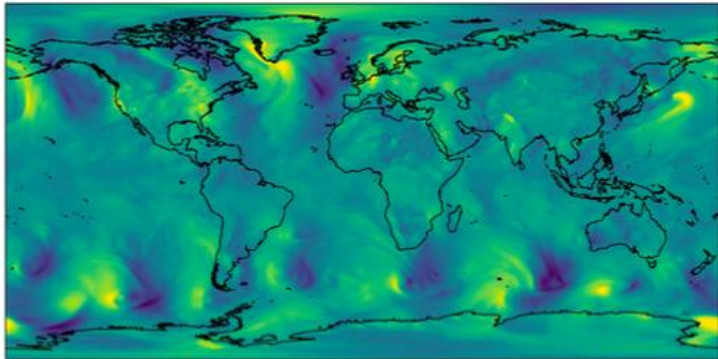
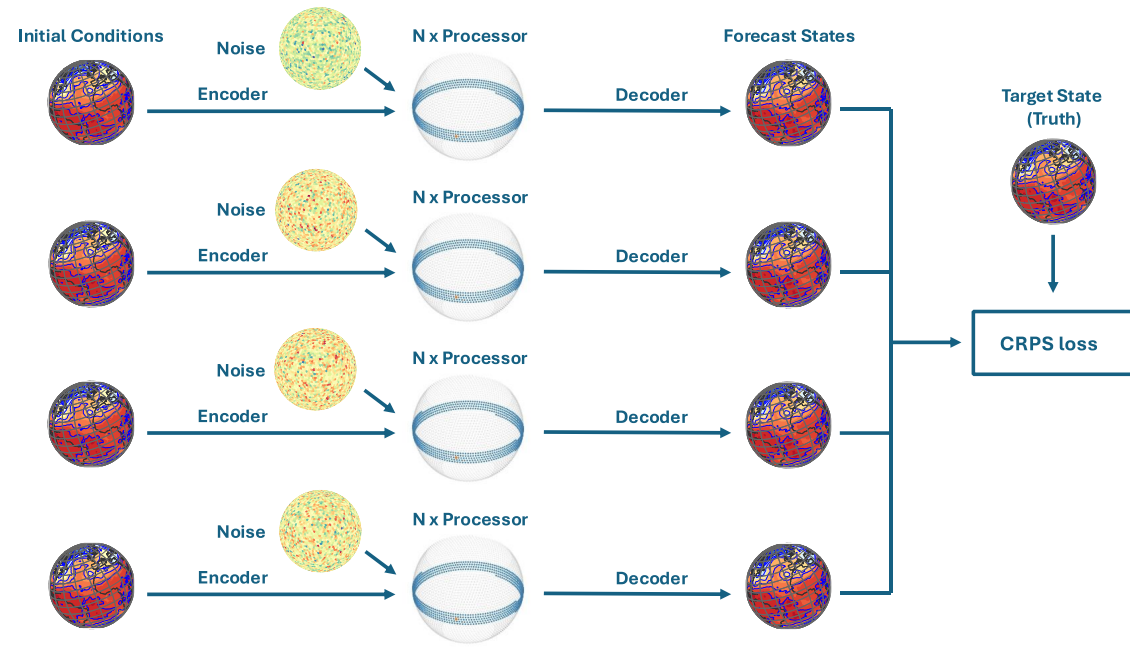
# AIFS-ENS:

Probabilistic training of AIFS:

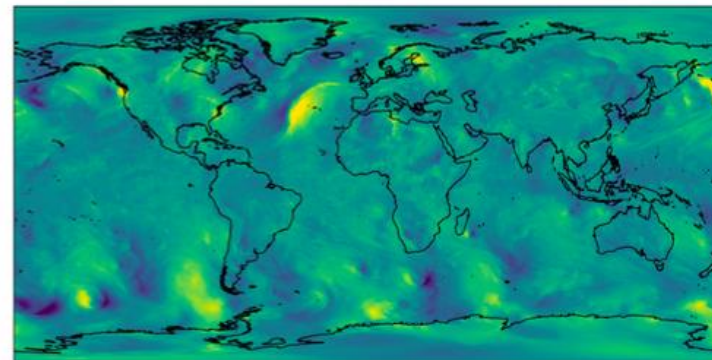
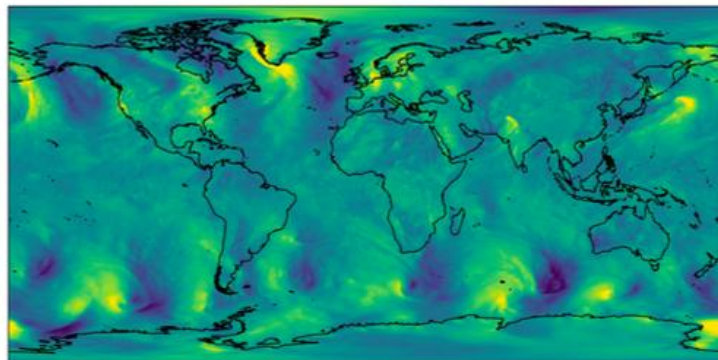
$$\text{afCRPS}_\alpha := \alpha \text{fCRPS} + (1 - \alpha) \text{CRPS}$$

$$= \frac{1}{M} \sum_{j=1}^M |x_j - y| - \frac{M-1+\alpha}{2M^2(M-1)} \sum_{j=1}^M \sum_{k=1}^M |x_j - x_k|$$

$$= \frac{1}{M} \sum_{j=1}^M |x_j - y| - \frac{1-\epsilon}{2M(M-1)} \sum_{j=1}^M \sum_{k=1}^M |x_j - x_k|$$



AIFS-single, day 1 and 10



AIFS-ENS, day 1 and 10

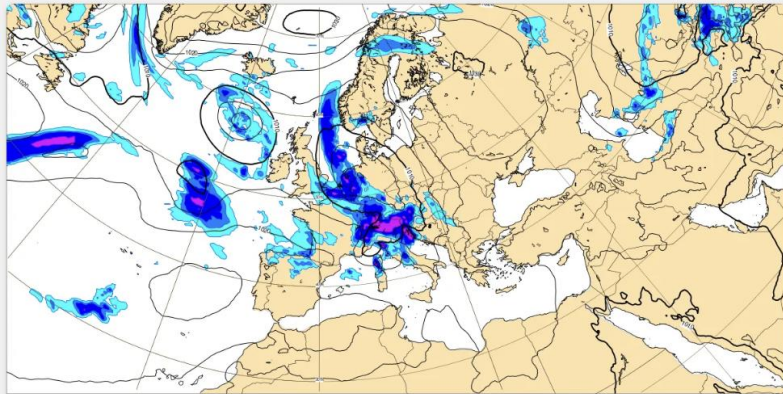
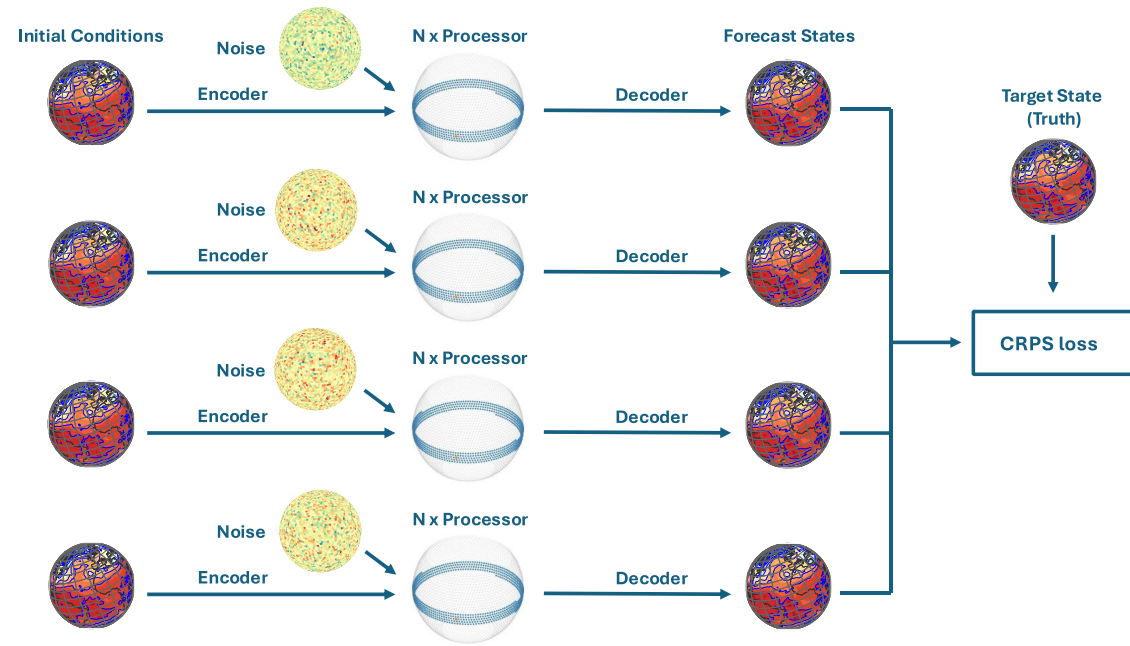
# AIFS-ENS:

## Probabilistic training of AIFS:

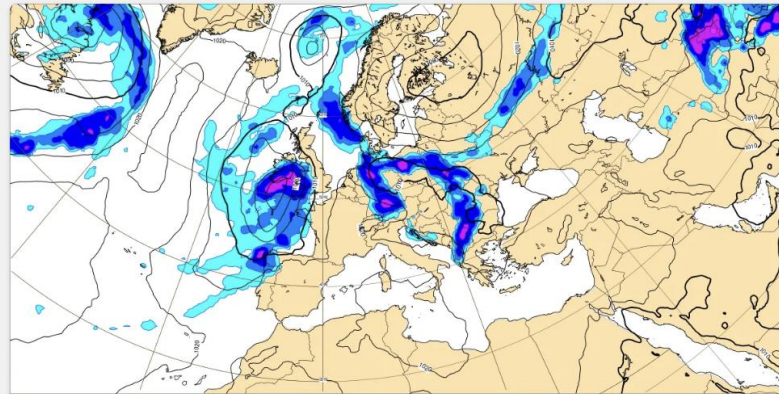
$$\text{afCRPS}_\alpha := \alpha \text{fCRPS} + (1 - \alpha) \text{CRPS}$$

$$= \frac{1}{M} \sum_{j=1}^M |x_j - y| - \frac{M-1+\alpha}{2M^2(M-1)} \sum_{j=1}^M \sum_{k=1}^M |x_j - x_k|$$

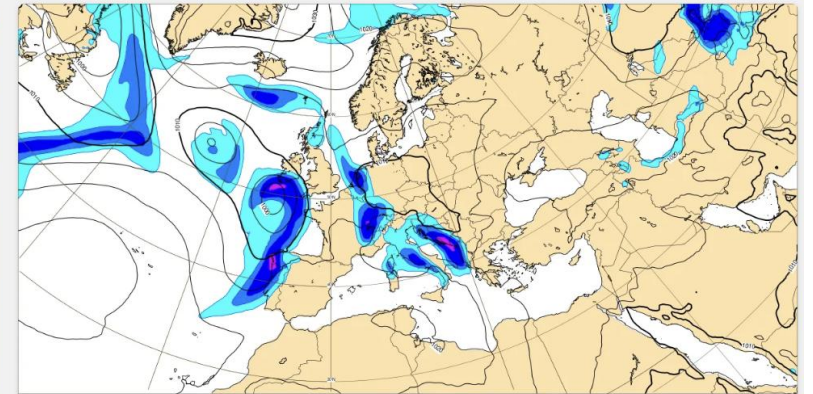
$$= \frac{1}{M} \sum_{j=1}^M |x_j - y| - \frac{1-\epsilon}{2M(M-1)} \sum_{j=1}^M \sum_{k=1}^M |x_j - x_k|$$



Rain and mean sea level pressure



AIFS CRPS : Rain and mean sea level pressure



AIFS Single: Rain and mean sea level pressure

10-day forecasts

# Running AIFS yourself interactively

**ecmwf/aifs-single-1.0** like 22 Following — ECMWF 56

Graph Machine Learning Anemol English doi:10.57967/hf/4629 arxiv:2406.01465 License: cc-by-4.0

Model card Files and versions Community Settings Use this model

### AIFS Single - v1.0

Here, we introduce the **Artificial Intelligence Forecasting System (AIFS)**, a data driven forecast model developed by the European Centre for Medium-Range Weather Forecasts (ECMWF).

The release of AIFS Single v1.0 marks the first operationally supported AIFS model. Version 1 supersedes the existing experimental version, [0.2.1 AIFS-single](#). The new version, 1.0, brings changes to the AIFS Single model, including among many others:

- Improved performance for upper-level atmospheric variables (AIFS Single still uses 13 pressure-levels, so this improvement mainly refers to 50 and 100 hPa)
- Improved skill for total precipitation.
- Additional output variables, including 100 meter winds, snow-fall, surface solar-radiation and land variables such as soil-moisture and soil-temperature.

Downloads last month **476**  
View full history

**Inference Providers** NEW

Graph Machine Learning

This model isn't deployed by any Inference Provider. Ask for provider support

**Collection including ecmwf/aifs-single-1.0**

**AIFS Single** Collection  
Artificial Intelligence Forecasting System · 4 items · Updated Feb 18



AIFS Single v1.1



AIFS Ens v1.0

- AIFS Single v1.0 and AIFS Ens v1.0 on Hugging Face – an open platform for sharing the models.
- Includes interactive notebook on how to run from ECMWF open data, can be adapted for other sources.

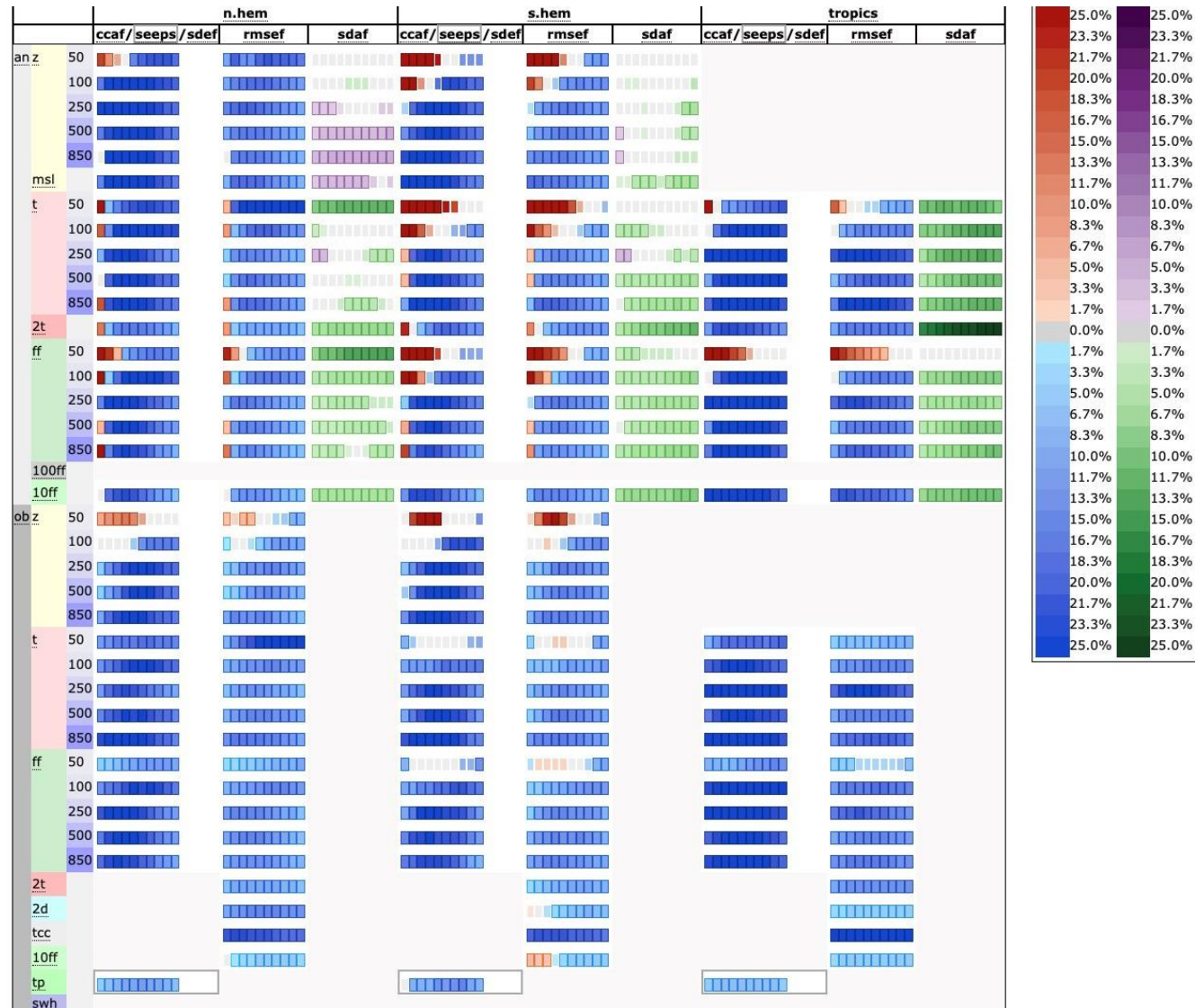
## 3. Load the Model and Run the Forecast

Download the Model's Checkpoint from Hugging Face & create a Runner

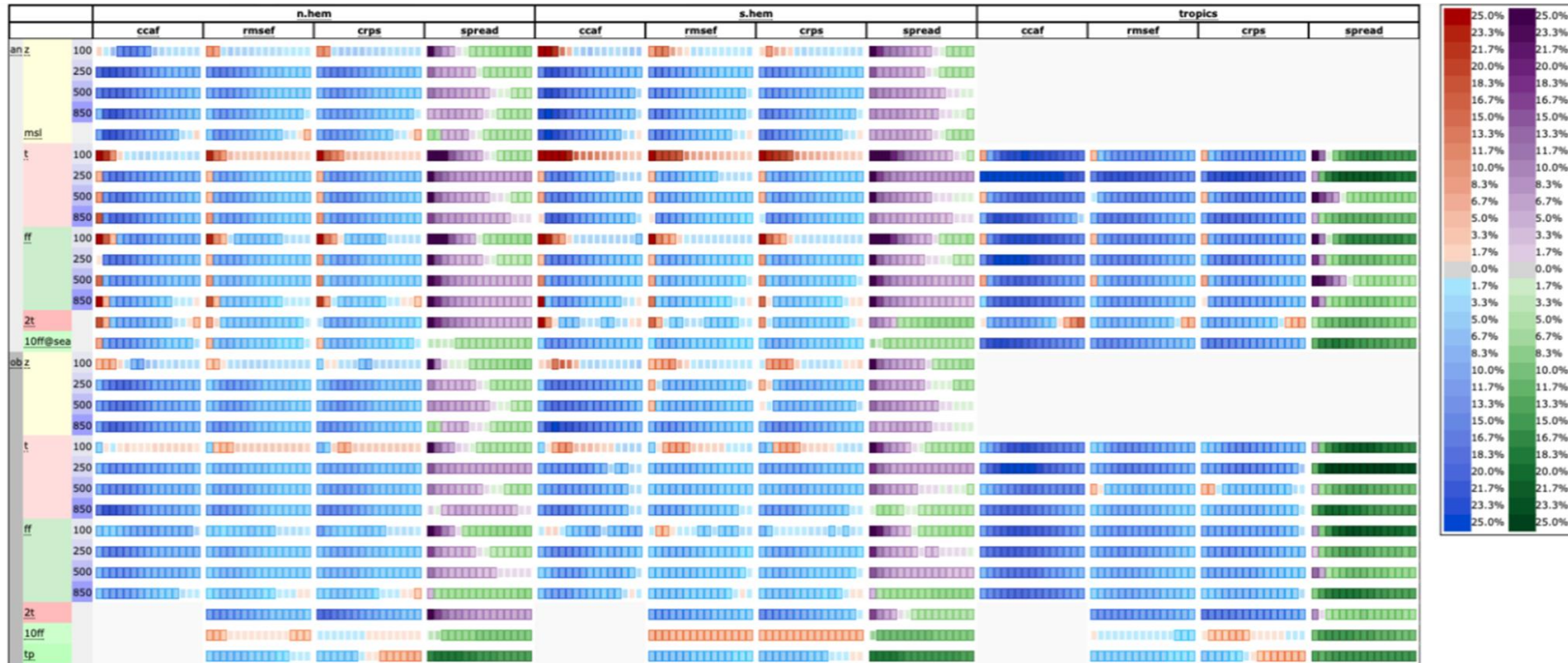
```
In [13]: checkpoint = {"huggingface": "ecmwf/aifs-single-1.0"}
```

Are these models skillful?

# AIFS Single v2 vs IFS



# AIFS ENS vs IFS



Are these models useful?

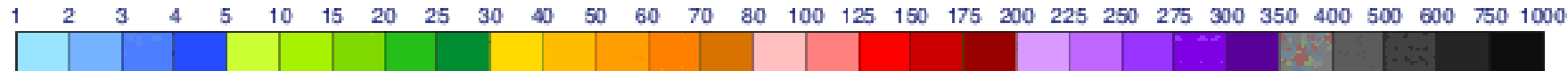
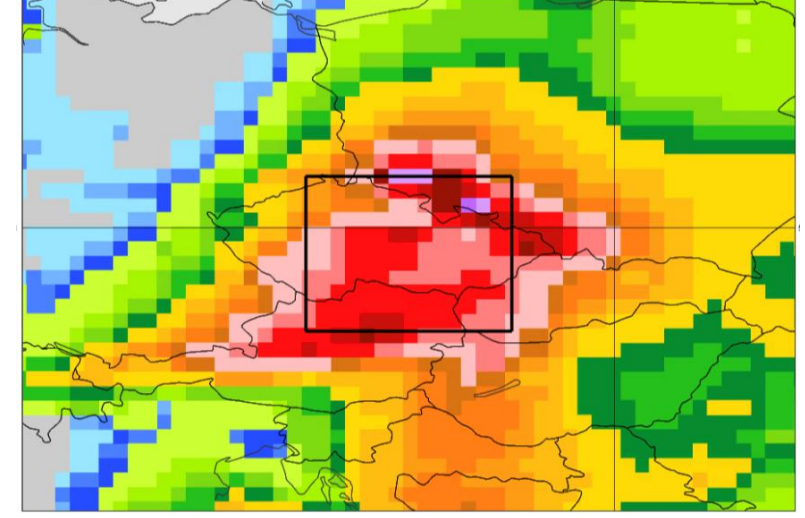
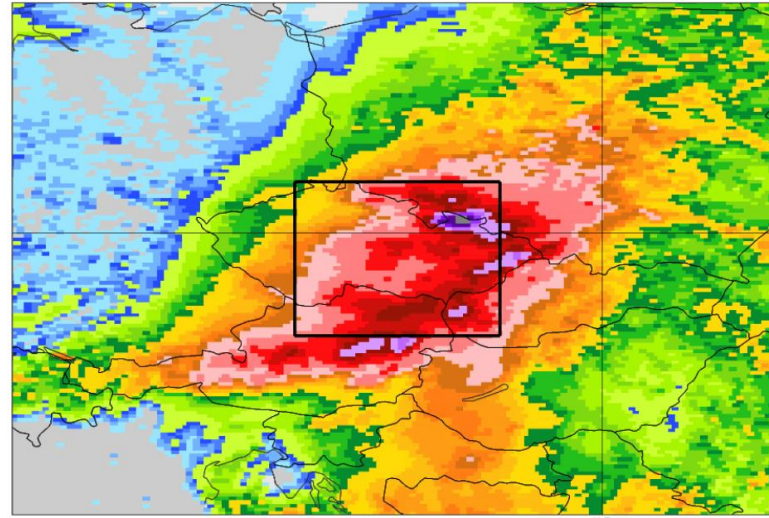
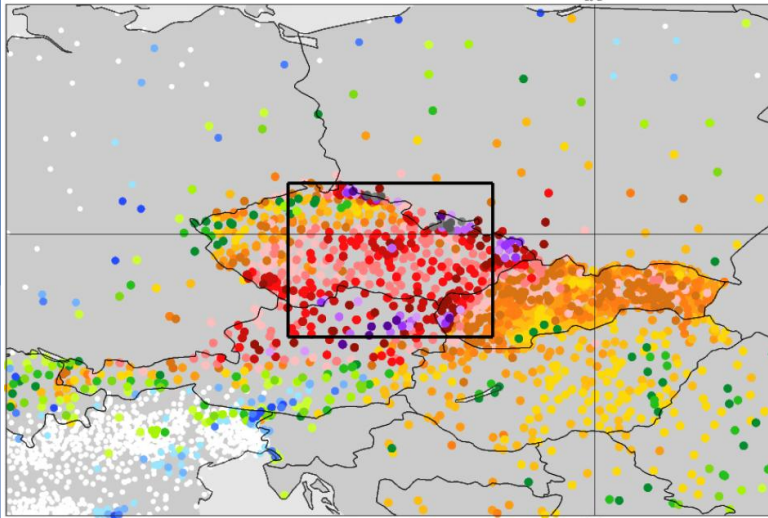
# Extreme precipitation during Storm Boris in central Europe, September 2024

72-hour precipitation 13 September 00UTC – 16 September 00UTC

Observations

IFS control (48-120h)

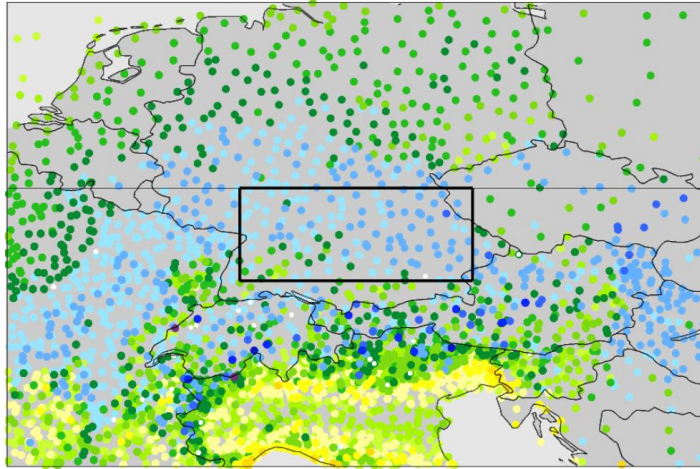
AIFSv1.0 (48-120h)



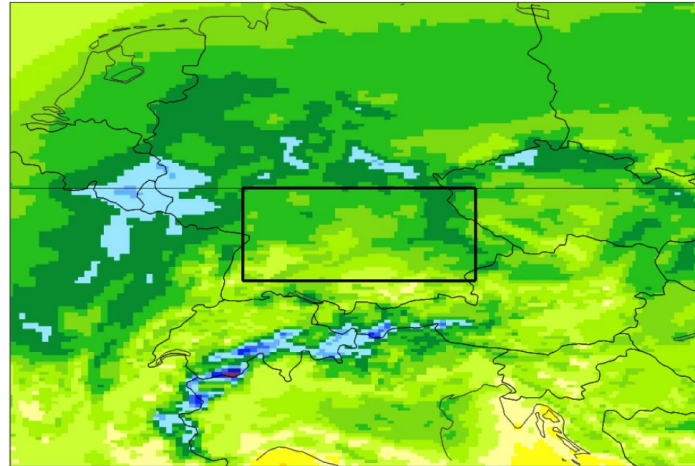
- AIFS predicted very extreme values for this region
- AIFS had a more consistent signal of the event over successive forecasts
- Smooth precipitation field from AIFS (do not capture local structures)

# Cold Event in Europe in December 2024

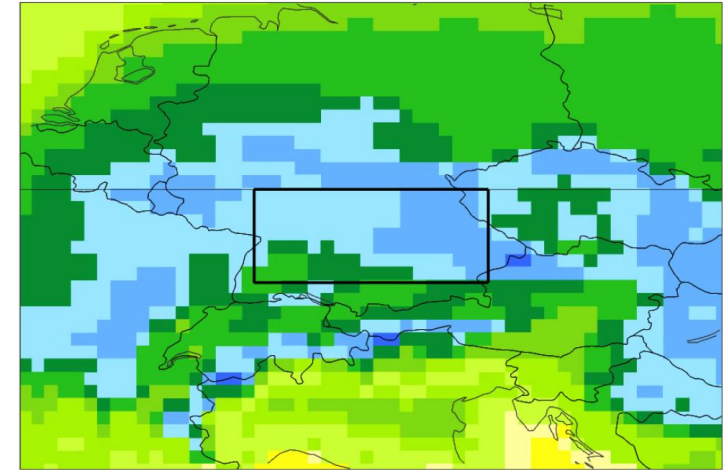
Observations



48-hour IFS Control



48-hour AIFS



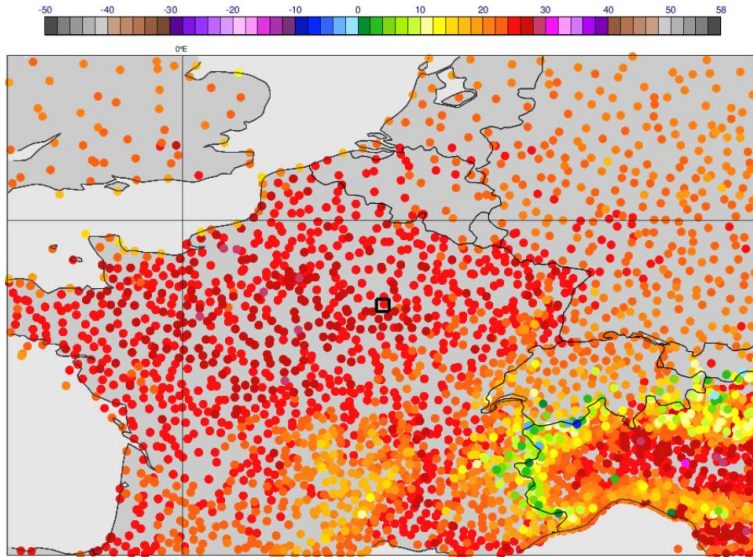
Karmouche et al. (2025):

AIFS blog: <https://www.ecmwf.int/en/about/media-centre/aifs-blog/2025/verifying-2-m-temperature-forecasts-wintertime-anticyclonic>

# Early heatwave in April 2025: 2-metre temperature 30 April 12UTC

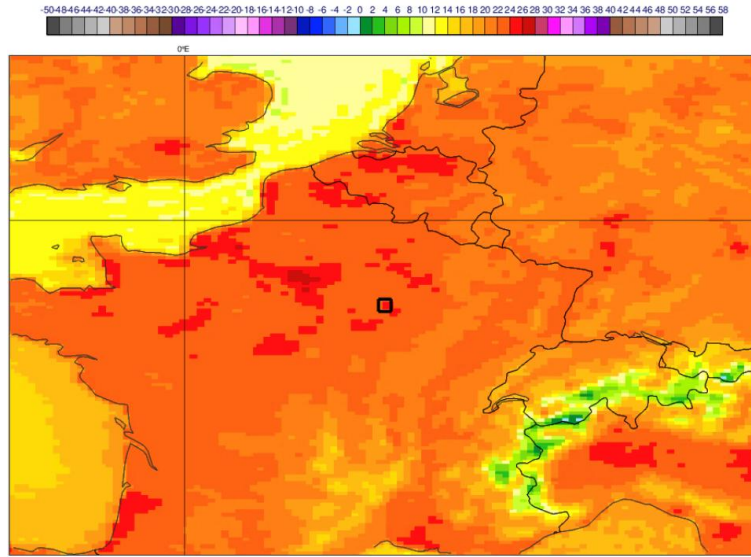
## Observations

2t\_Troyes0.0\_2025043012\_od\_oper\_ecmf\_0001\_obs  
Observations



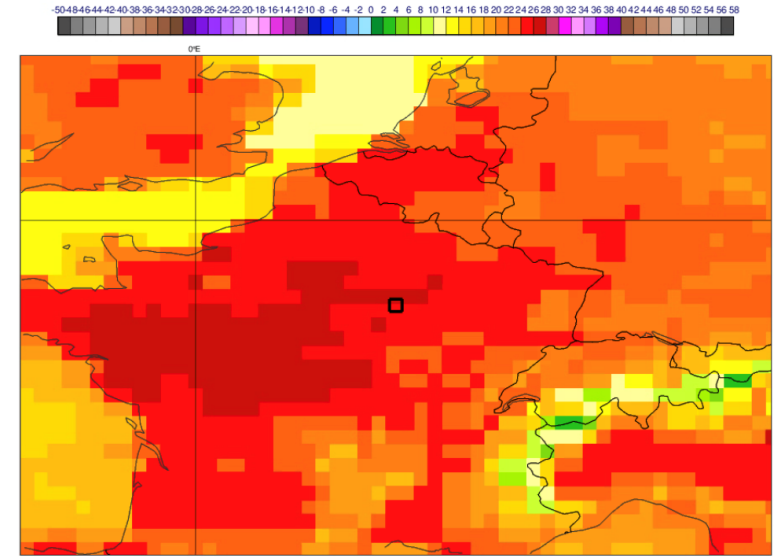
## 24h forecast from IFS-ENS Mem 1

2t\_Troyes0.0\_2025043012\_od\_enfo\_ecmf\_0001\_pf  
Member: 1, Step: 24



## 24h forecast from AIFS-ENS Mem 1

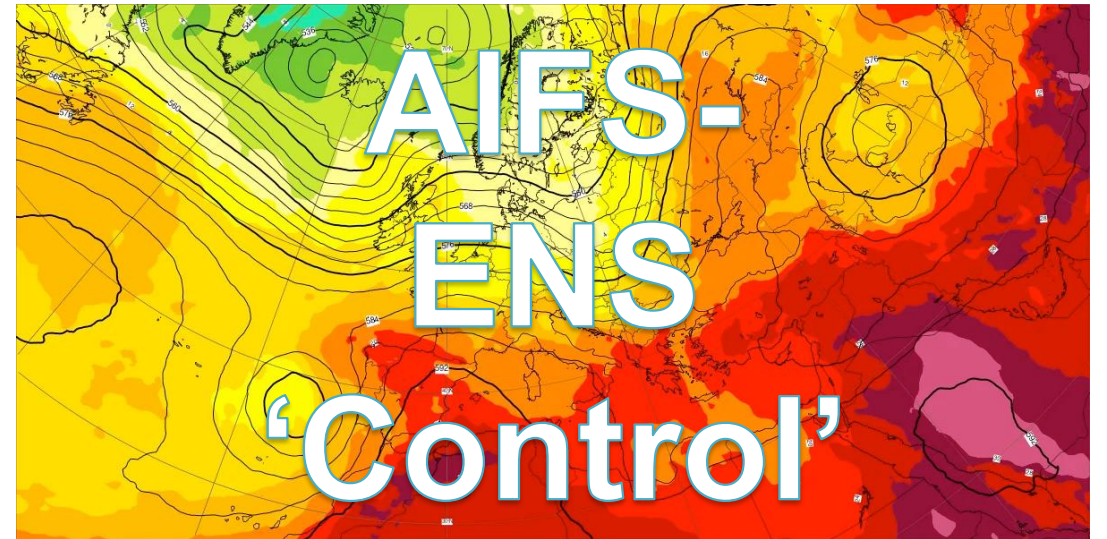
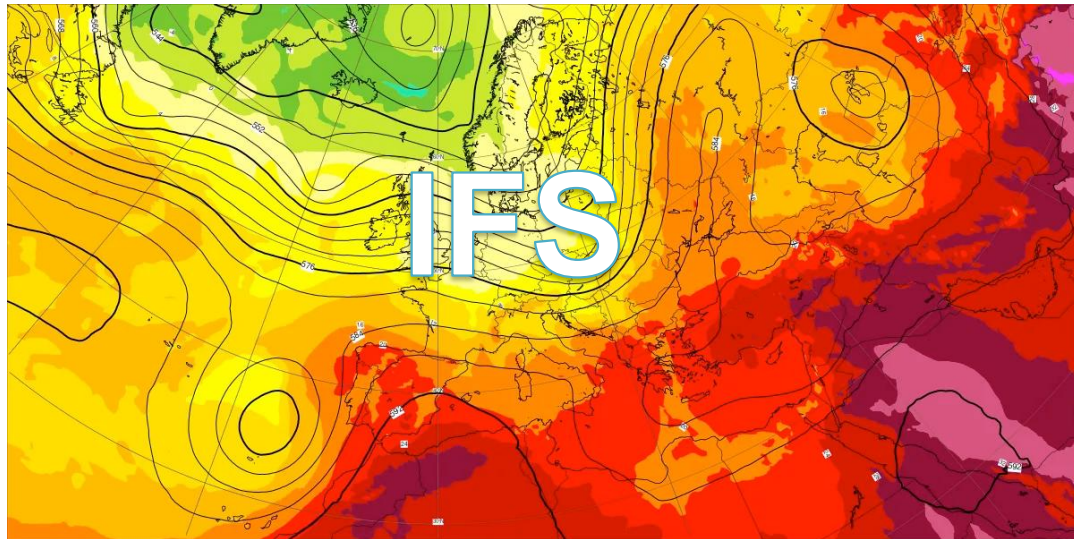
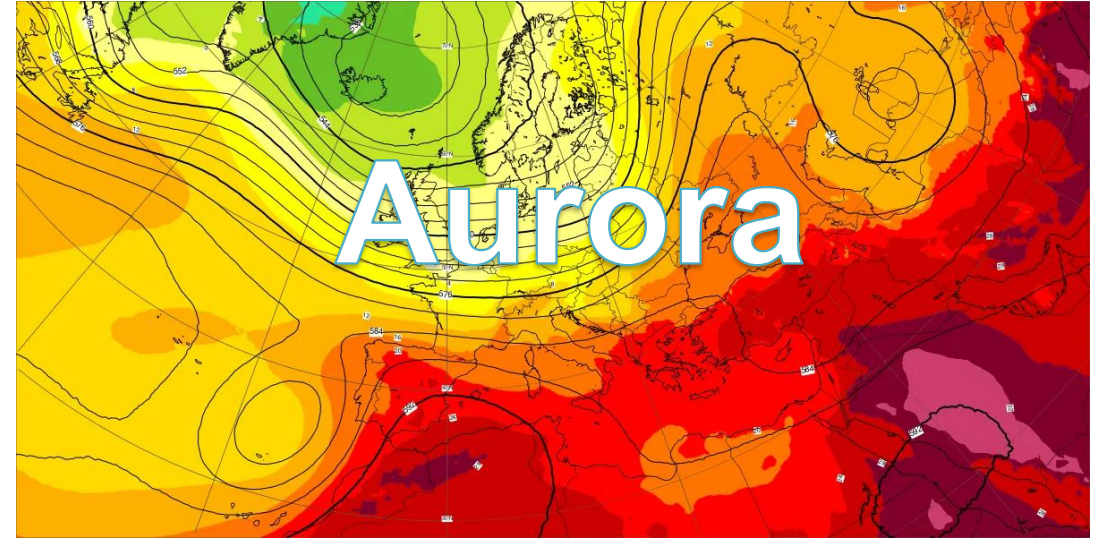
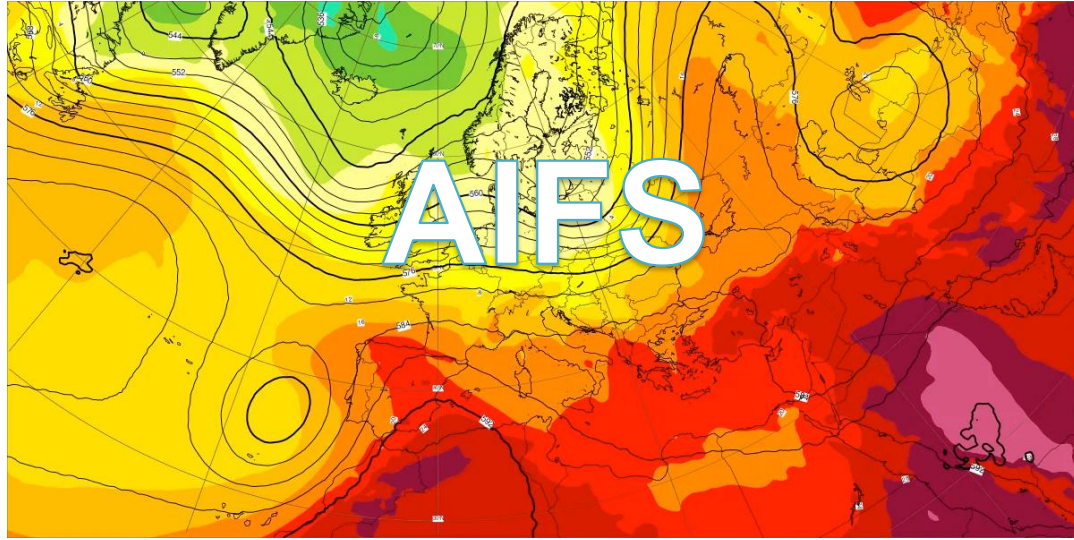
2t\_Troyes0.0\_2025043012\_ai\_enfo\_ecmf\_9101\_pf  
Member: 1, Step: 24



T+156

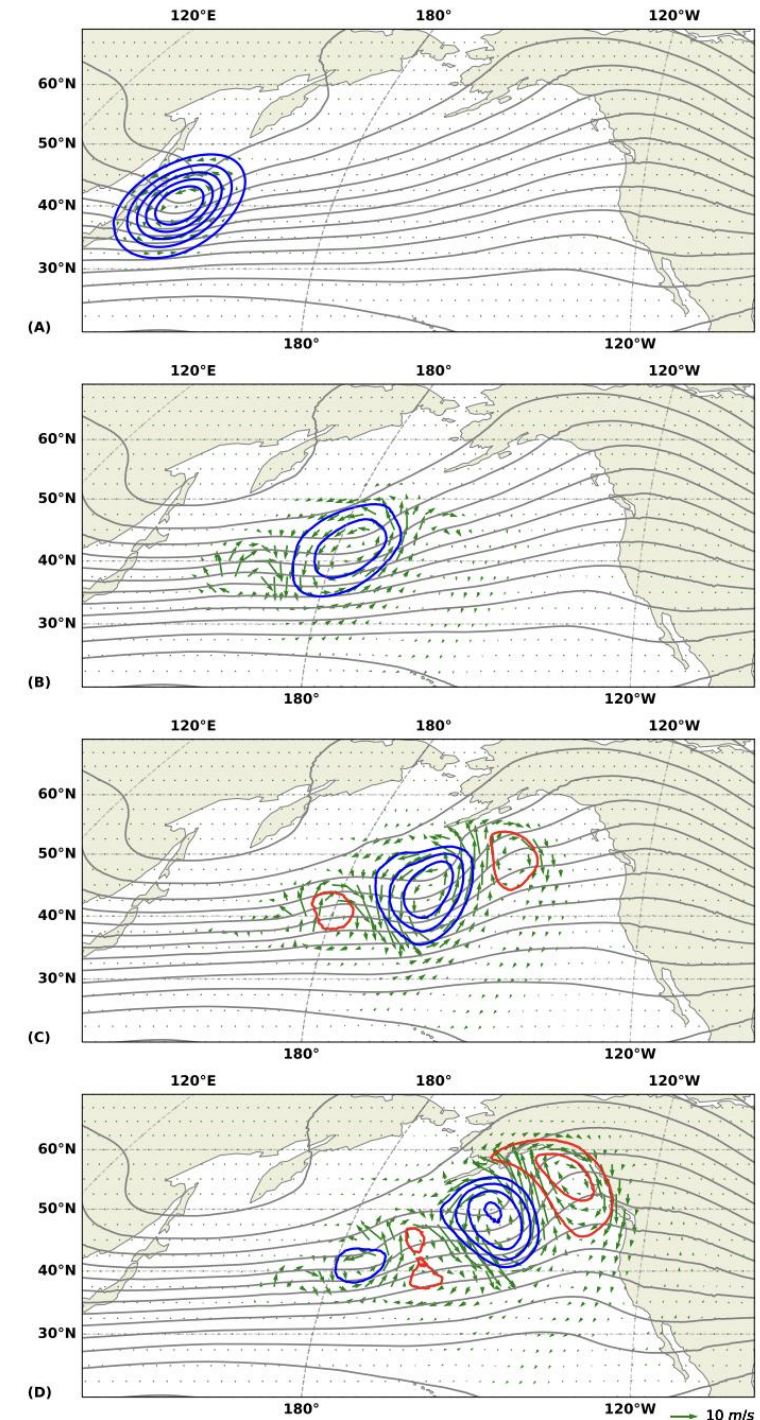
Basetime: Tue 03 00 UTC

Validtime: Mon 09 12 UTC



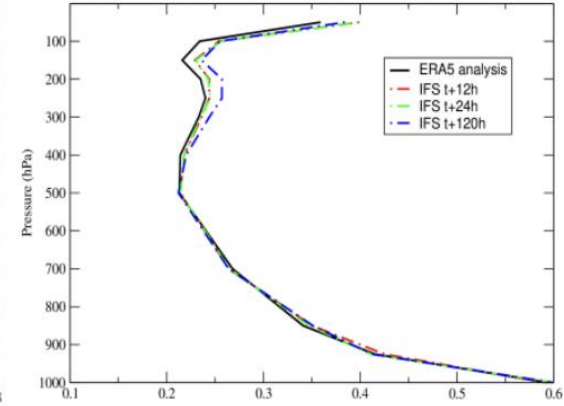
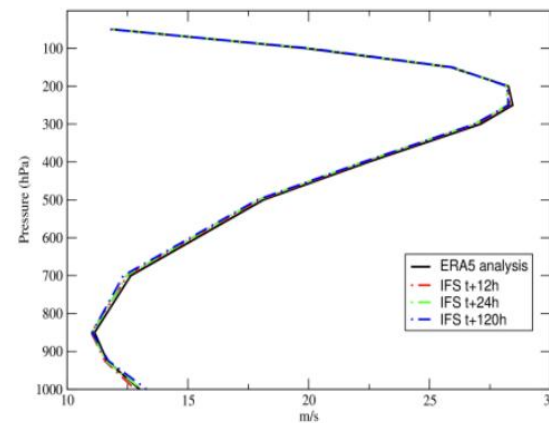
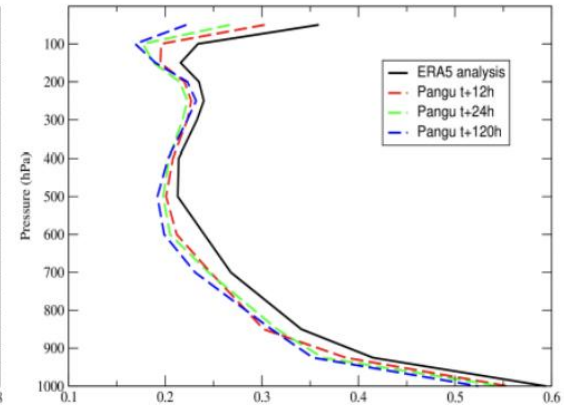
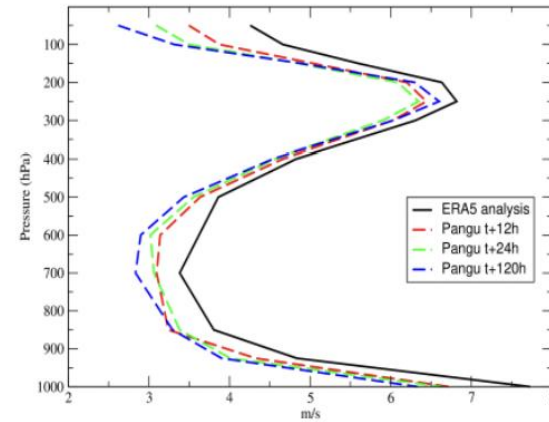
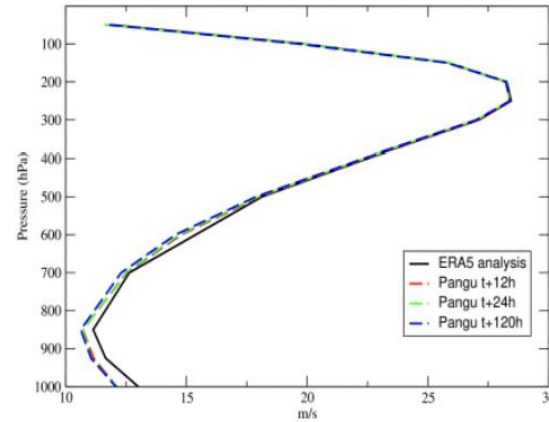
# Are data-driven weather forecasts physical?

- Highly recommend reading Hakim & Masanam 2023:  
Dynamical Tests of a Deep-Learning Weather Prediction Model
- Take Pangu weather, and test it on a series of classical dynamical core test cases.
  - Cases need to be applied as deviations from climatology.
  - Apply localised disturbances and study the reaction of the system.
- Overall, Pangu behaves as expected, compared with theory.
  - The 1h model is best for the faster evolving processes, which aren't well captured by the longer timestep model.
- Hopefully we see lots more studies of this type.



# Are data-driven weather forecasts physical?

- Bonavita 2023 explore other tests with Pangu Weather.
- Geostrophic balance fairly well represented, but not as well as the IFS.



Geostrophic winds

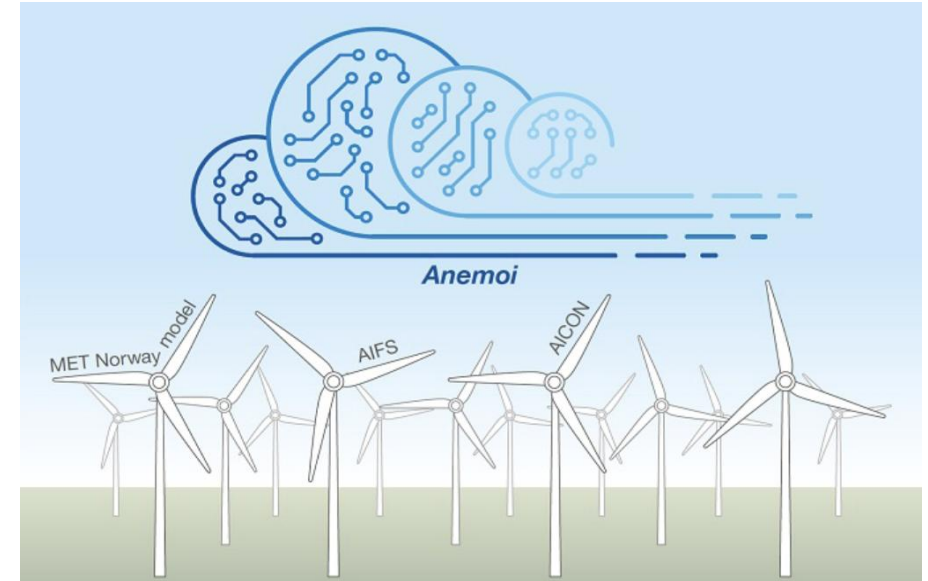
Ageostrophic winds

Ratio

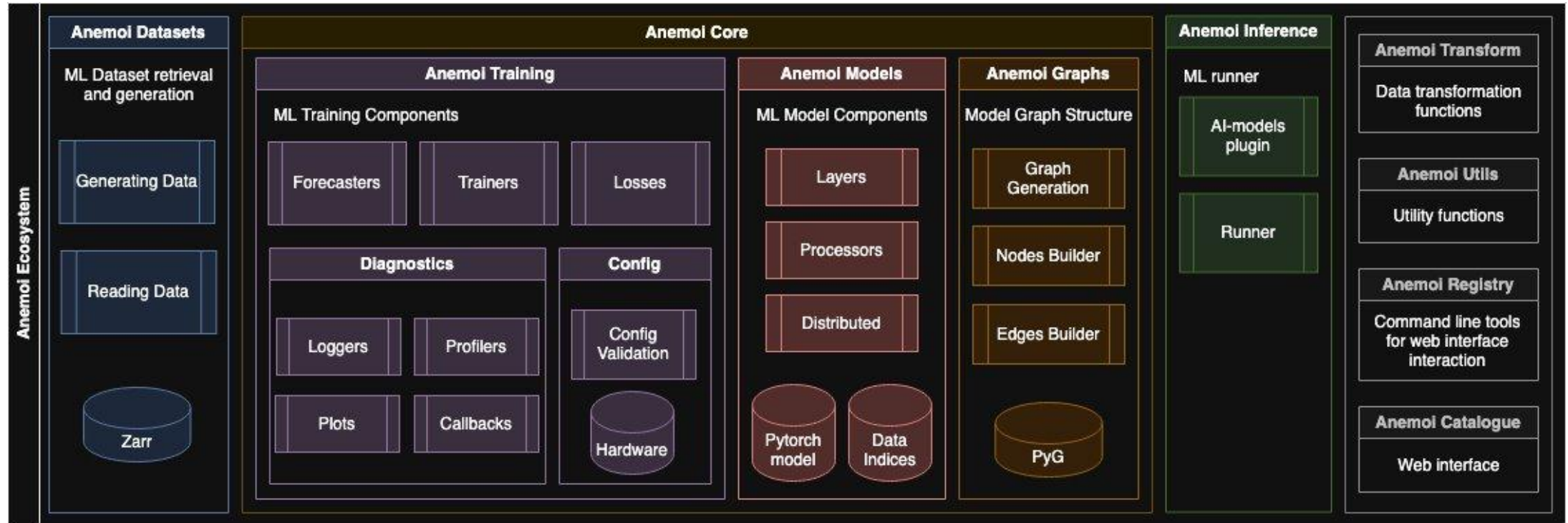
# Anemoi: the tool behind it all

# What is Anemoi?

- An *open-source framework* that provides a complete toolkit to develop data-driven weather models – from data preparation through to inference
- Empowering *real-time running* and operationalisation of data-driven models at meteorological centres
- A *highly modular framework* organised into different Python packages



# What is Anemoi?



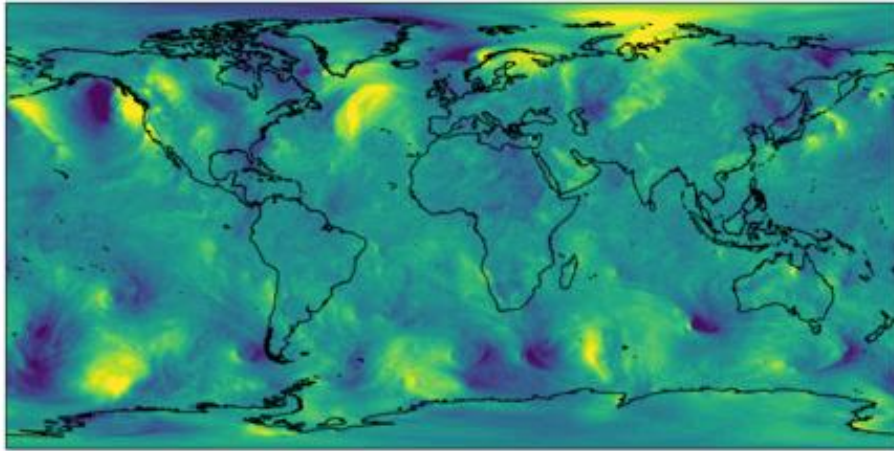
# Who is Anemoi?

Pooling of resources without resulting in a single forecasting model.



Open-source development driven by European Meteorological Centers and ECMWF

# What Anemoi enables?

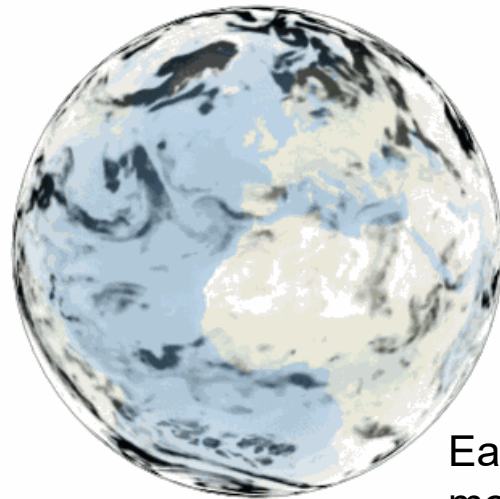
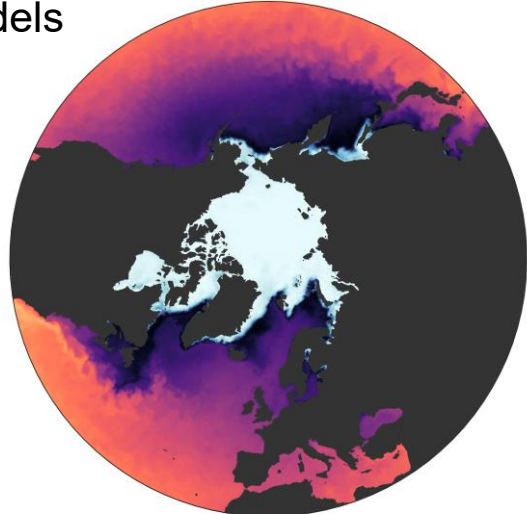


Training deterministic & ensemble models

Lang et al. 2024a & b

## Building coupled earth-system models

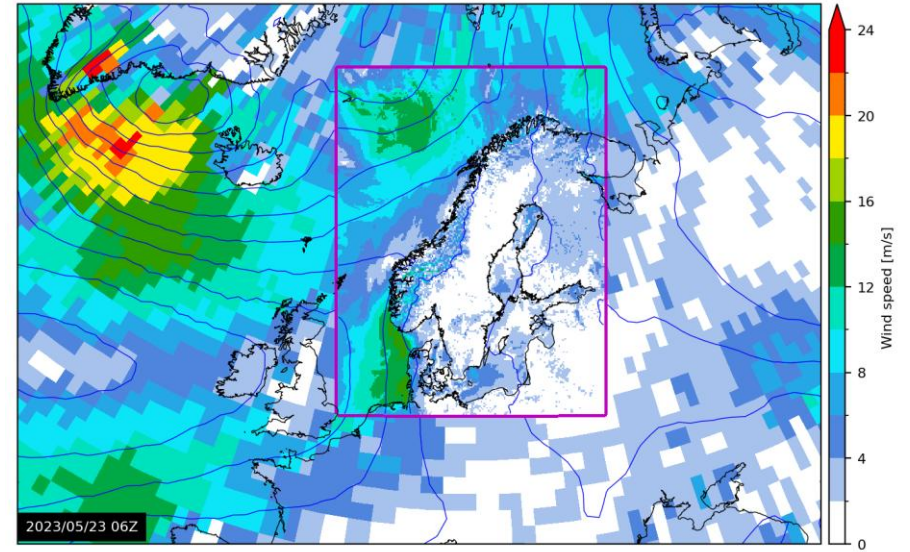
ARCTIC - 2022-01-03



Easy exploration of more atmospheric fields

Regional high-resolution modelling:

Nipen et al. 2024



Real-time running at operation centres  
Anemoi is powering ECMWF's operationalisation of AIFS in 2025.

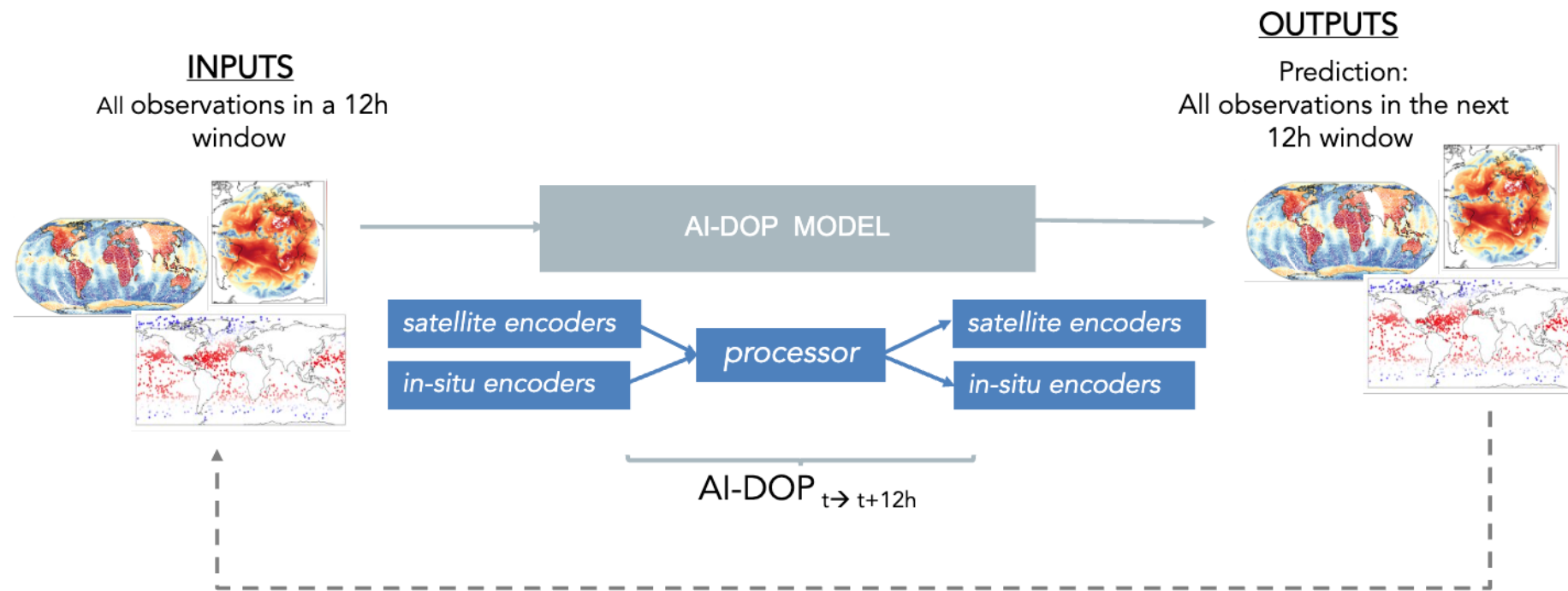
A screenshot of a web interface showing three forecast panels. The first panel is titled 'AIFS CRPS Control: Mean sea level pressure and 850 hPa wind speed'. The second panel is titled 'AIFS Single: Mean sea level pressure and 850 hPa wind speed'. The third panel is titled 'AIFS CRPS Control: 500 hPa geopotential height and 850 hPa temperature'. Each panel includes a map, a 'Latest forecast' section, and a 'Wind speeds near the surface are roughly proportional to the distance between isobars so closely packed isobars mean strong surface winds...' note. The interface also features navigation icons for 'jupyter', 'github', 'launch binder', and 'Open in Colab'.

# Learning directly from observations

# AI-DOP: Emulating the entire chain using only observations

Alexe et al. 2024

- Use historical observations to train a Neural Network to forecast future observations (don't need analysis)
- Include all available observations of the full Earth system (atmosphere, ocean, land) simultaneously

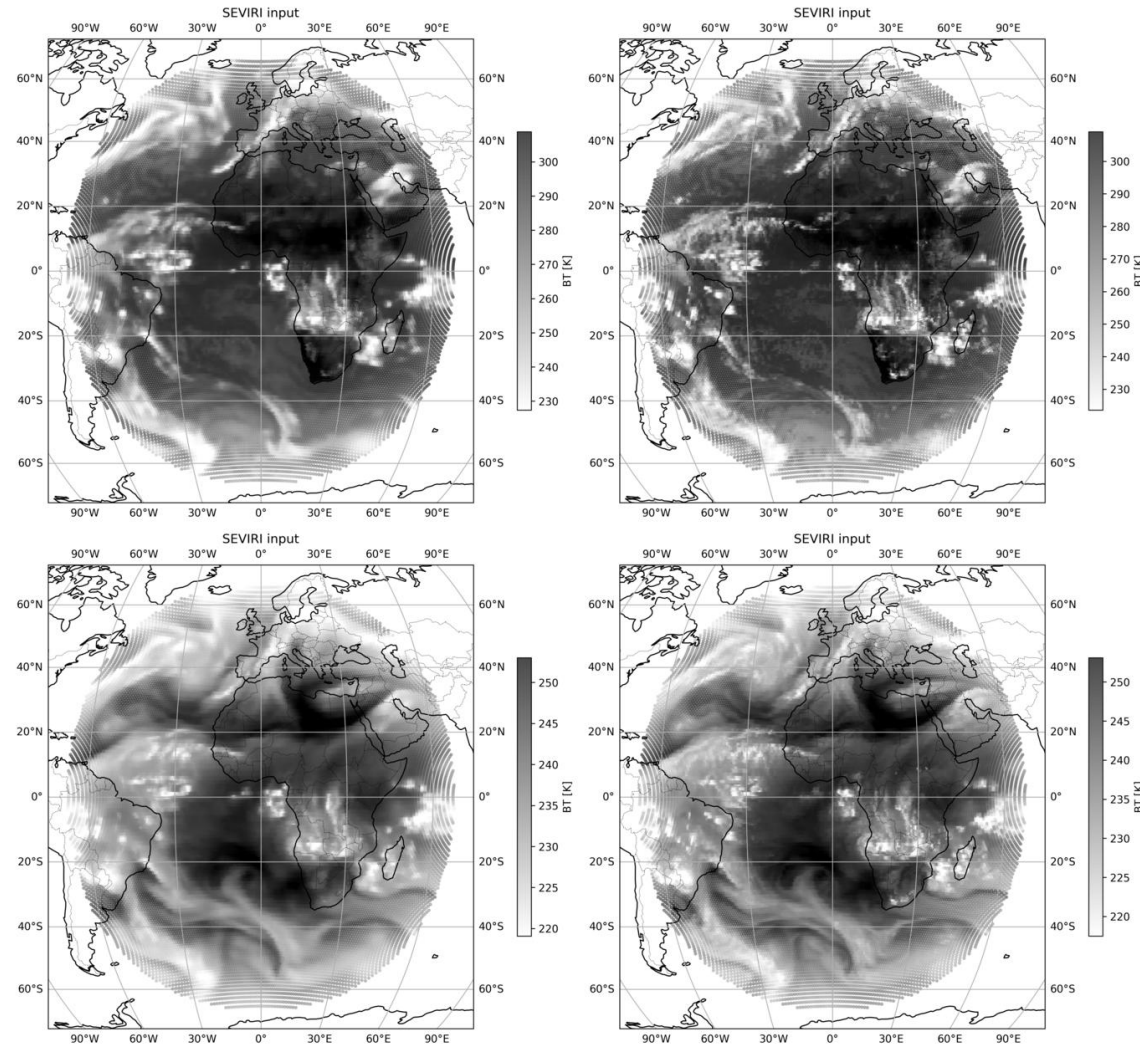


# AI-DOP: Emulating the entire chain using only observations

Alexe et al. 2024

SEVIRI prediction

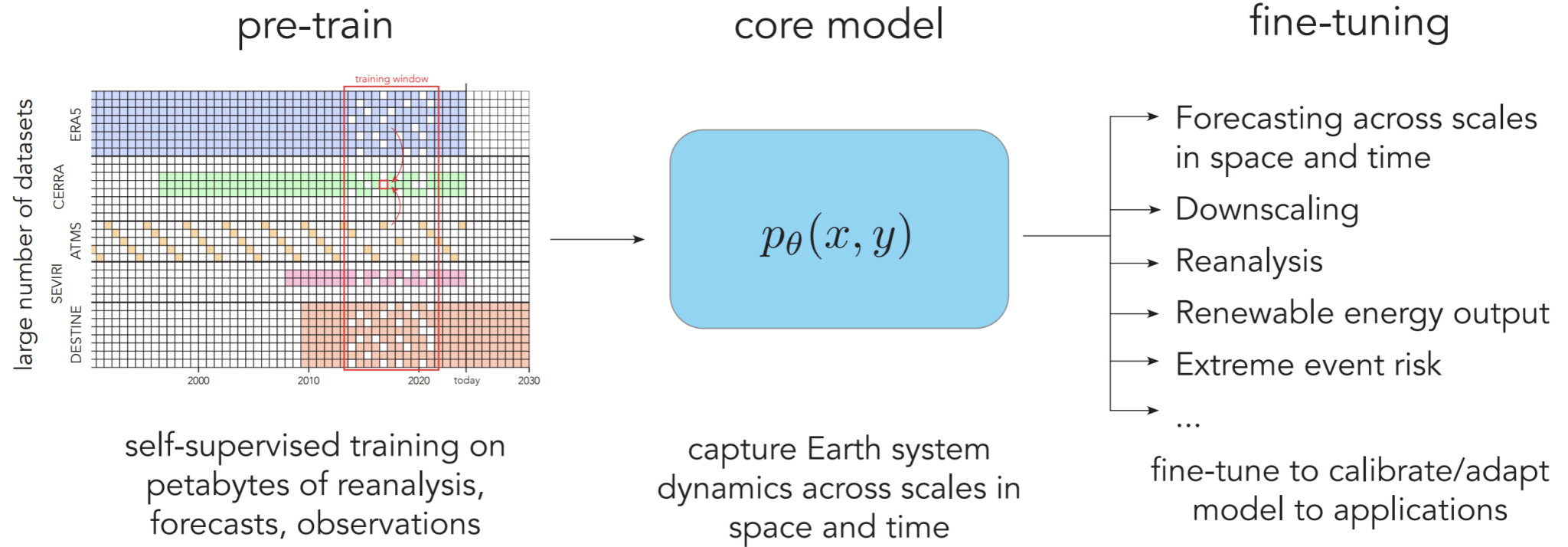
SEVIRI target



# WeatherGenerator



A pan-European initiative to build a single open-source foundational model for weather and climate



# Summary

- Data-driven weather forecasting is a fast-evolving and exciting landscape
- With more advances to come:
  - Full Earth System models
  - Direct Observation Prediction models
  - Foundational models
  - ....